The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
Abstract

Domain Generality and Individual Differences in Disjunctive Thinking

Doctor of Philosophy

2001

Maggie E. Toplak

Department of Human Development and Applied Psychology

OISE/University of Toronto

Disjunctive thinking is an essential and ubiquitous critical thinking skill. The framework used to examine disjunctive thinking was taken from Shafir (1994), who termed disjunctive thinking as the tendency to consider all of the possible states of the world and to evaluate the consequences of each. Shafir (1994) argued that people are typically reluctant to think through disjunctions, and a failure to reason disjunctively is a violation of one of the most basic principles of rational choice, called Savage’s Sure Thing Principle. Data from one hundred and twenty-five participants (47 males and 78 females) was examined on a set of formal and informal tasks, taken from the reasoning and decision-making literatures. It was found that 21% to 86% of participants failed to reason disjunctively on the formal tasks. Evidence for a failure to reason disjunctively was also found on one informal reasoning task, called The Argument Generation Task. There was some small support for the domain generality of a disjunctive thinking skill in the formal reasoning problems, and for a generational fluency in the informal tasks. However, the formal reasoning, formal decision-making, and informal tasks were virtually unrelated, providing support for domain specificity in the problems used here to investigate disjunctive thinking. The individual difference measures of
cognitive ability and thinking dispositions were found to predict performance on some of the formal tasks. In particular, the dispositional measures: the Need for Cognition Scale and errors on the Matching Familiar Figures Test, were found to explain unique variance on some of the formal problems. These findings provide support for malleable, dispositional explanations of performance on the formal tasks. The implications of these findings are discussed in terms of the Generic Dual Process Framework (Stanovich, 1999; Stanovich & West, 2000).
Acknowledgements

I would like to acknowledge and thank my committee members, Keith Oatley, Uri Shafrir, and Richard F. West. Their interest and engagement with my research has been helpful throughout this project. In particular, their participation in my oral examination was a highlight of my intellectual experience. A very special thank you is extended to my supervisor and mentor Keith Stanovich. He has equipped me with all of the necessary skills to be a successful scientist, intellectual, professional, and person. His mentoring has provided me with more than I could ever give back to him. His love for science and investigation is contagious! And Keith, this is only the beginning!

The members of the Stanovich Lab have also been a source of continual support, with constant sharing of ideas and encouragement. Walter C. Sá provided helpful comments on my thesis, and he has also been a good friend.

My family has also been very supportive since the beginning. My parents, Vera Toplak (nee Feher) and Joseph Toplak, have waited patiently and in anticipation. They taught me very early on to appreciate and to value education, because education was something that no one could take away from you, and they are right! My brother Tim Toplak helped me with some data entry related to this project. He has been eager and willing to help throughout.

- Thanks Tim!

Of course Paul H. Voorn has been the rock, the dishwasher, and the loudest pep rally! When I was working late at school, you would bring me shrimps in sour soup. When I was working late at home, you would bring me tea. Without any hesitation, you have shared everything with me. Paul, you are my best buddy, and I could not have done this without you by my side!
# Table of Contents

Abstract .................................................................................................................. ii
Acknowledgements ................................................................................................. iv
Table of Contents ..................................................................................................... v
List of Tables ............................................................................................................. x
List of Figures .......................................................................................................... xiv
List of Appendices .................................................................................................... xv

## Introduction
The Domain Generality of Disjunctive Thinking: A Taxonomy of Tasks ............. 2
Disjunctive Consideration in Decision-Making Problems .................................. 3
Disjunctive Consideration in Reasoning Problems ............................................. 9
Logic of the Disjunctive Tasks Used in this Study .............................................. 13
What Factors are Associated with Individual Differences in Disjunctive Thinking? The Distinction Between Cognitive Capacities and Thinking Dispositions ................................................................. 14
Thinking Dispositions Examined in the Present Investigation ......................... 23
   The Dimension of Reflectivity and Impulsivity ........................................... 23
   The Need For Cognition .............................................................................. 25
   The Actively Open-Minded Thinking Disposition ................................... 28
Summary ............................................................................................................... 28

## Method
Participants ............................................................................................................. 30
Outliers .................................................................................................................. 30
Decision-Making and Reasoning Tasks ............................................................... 33
   Formal Decision Making Tasks .................................................................. 33
      The Prisoner’s Dilemma ................................................................. 33
      Newcomb’s Problem ........................................................................... 35
      The Disease Framing Problem ....................................................... 36
      The Box Problem ................................................................................ 38
   Informal Decision Making Task ............................................................... 39
      The Consumer Task .......................................................................... 39
Formal Reasoning Problems ................................................................. 40
The Selection Task ............................................................................ 40
Knights and Knaves Problem ............................................................... 42
Charles in Scotland Problem ............................................................... 43
The Married Problem ......................................................................... 44
The Green Levels Problem ................................................................. 45
Informal Reasoning Task ................................................................. 46
The Argument Generation Task ......................................................... 46
The Argument Generation Task and the Role of Knowledge .......... 49
Thinking Dispositions ......................................................................... 51
The Matching Familiar Figures Test (MFFT) .................................... 51
Thinking Dispositions Questionnaire .............................................. 52
The Need for Cognition Composite ................................................... 52
The Actively Open-Minded Thinking (AOT) Composite ............... 53
Other Dispositions ............................................................................ 54
General Ability Measures ................................................................... 55
Procedure .......................................................................................... 57

Results and Discussion
1. Description of Performance .............................................................. 58
Formal Decision Making Problems .................................................. 58
Prisoner’s Dilemma ............................................................................ 58
Newcomb’s Problem ......................................................................... 60
Disease Framing Problem .................................................................. 60
Box Problem ....................................................................................... 61
Informal Decision-Making Task ......................................................... 62
The Consumer Task .......................................................................... 62
Formal Reasoning Problems ................................................................ 63
The Selection Task ............................................................................ 63
Knights and Knaves Problem ............................................................... 66
Charles in Scotland Problem ............................................................... 66
Married Problem ................................................................................ 69
Green Levels Problem ......................................................................... 70
Informal Reasoning Task .................................................................... 70
2. Domain Generality ................................................................. 72
   Formal Decision-Making and Formal Reasoning Problems .......... 72
   Knights and Knaves, Married Problem, Green Levels Problem .......... 72
   Charles in Scotland Problem .................................................. 75
   Other Comparisons with the Charles in Scotland Problem .......... 76
   Selection Task ........................................................................ 78
   Selection Task: P Choosers versus P, Q Choosers ....................... 78
   Selection Task: All Choosers versus P, Q Choosers ....................... 79
   Box Problem, Framing Disease Problem, Newcomb, Prisoner's
   Dilemma .................................................................................. 80
   Informal Tasks ......................................................................... 81
   Consumer Task ......................................................................... 81
   Argument Generation Task ....................................................... 82
   Consumer Task and Argument Generation Task ........................... 87
   Formal and Informal Tasks .......................................................... 88
   The Consumer Task and Formal Decision-Making and Reasoning
   Problems .................................................................................. 89
   The Argument Generation Task and Formal Decision-Making and
   Reasoning Problems .................................................................. 89

3. Predicting Performance on Formal and Informal Tasks using Cognitive
   Ability and Thinking Dispositions ............................................. 90
   Formal Decision-Making Problems ............................................ 91
   Prisoner's Dilemma ................................................................. 91
   Newcomb's Problem ............................................................... 95
   Disease Framing Problem ......................................................... 95
   Box Problem ........................................................................... 98
   Formal Reasoning Problems .................................................... 101
   Selection Task ........................................................................ 101
   Selection Task: P Choosers versus P, Q Choosers ....................... 102
   Knights and Knaves Problem .................................................... 103
   Charles in Scotland Problem .................................................... 103
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Married Problem</td>
<td>107</td>
</tr>
<tr>
<td>Green Levels Problem</td>
<td>109</td>
</tr>
<tr>
<td>Summary of Individual Difference Measures and Regression Analyses</td>
<td>111</td>
</tr>
<tr>
<td>Informal Decision-Making Task</td>
<td>113</td>
</tr>
<tr>
<td>The Consumer Task</td>
<td>113</td>
</tr>
<tr>
<td>The Consumer Task and Prior Experience</td>
<td>120</td>
</tr>
<tr>
<td>Informal Reasoning Task</td>
<td>120</td>
</tr>
<tr>
<td>Argument Generation Task</td>
<td>120</td>
</tr>
<tr>
<td>Argument Generation Task and Prior Belief</td>
<td>120</td>
</tr>
<tr>
<td>Argument Generation Task and Knowledge</td>
<td>124</td>
</tr>
<tr>
<td>Other Notable Findings</td>
<td>126</td>
</tr>
<tr>
<td>Thinking Dispositions</td>
<td>126</td>
</tr>
<tr>
<td>Concluding Discussion</td>
<td>128</td>
</tr>
<tr>
<td>Is Disjunctive Thinking A Domain General Skill?</td>
<td>129</td>
</tr>
<tr>
<td>Patterns of Associations on the Formal Reasoning Tasks</td>
<td>129</td>
</tr>
<tr>
<td>What do we make of Newcomb's Problem and the Prisoner's Dilemma?</td>
<td>131</td>
</tr>
<tr>
<td>Newcomb</td>
<td>131</td>
</tr>
<tr>
<td>Prisoner's Dilemma</td>
<td>134</td>
</tr>
<tr>
<td>Patterns of Associations on the Informal Reasoning Tasks</td>
<td>137</td>
</tr>
<tr>
<td>Evidence for 'Generational Fluency.'</td>
<td>137</td>
</tr>
<tr>
<td>Where There Was No Generational Fluency</td>
<td>138</td>
</tr>
<tr>
<td>Summary: Domain Generality and Disjunctive Thinking in Informal and Formal Tasks</td>
<td>140</td>
</tr>
<tr>
<td>Who Does Well on Disjunctive Thinking Problems?</td>
<td>141</td>
</tr>
<tr>
<td>Formal Reasoning and Decision-Making Problems</td>
<td>141</td>
</tr>
<tr>
<td>The Special Case of P Only Choosers in the Selection Task</td>
<td>145</td>
</tr>
<tr>
<td>Explaining Performance on the Informal Tasks</td>
<td>147</td>
</tr>
<tr>
<td>Why do People Fail to Reason Disjunctively?: Implications for Models of Human Reasoning</td>
<td>148</td>
</tr>
<tr>
<td>Limitations of the Present Study</td>
<td>152</td>
</tr>
</tbody>
</table>
List of Tables

Table 1: Decision Making Tasks and Reasoning Problems from Shafir (1994; Shafir & Tversky, 1992)........................................................................................................4
Table 2: Formal and Informal Reasoning and Decision-Making Tasks Used in This Study...........................................................................................................13
Table 3: The Terms for the Two Systems Used by a Variety of Theorists and the Properties of Dual-Process Theories of Reasoning (reproduced from Stanovich, 1999)...........................................................................................................19
Table 4: Formal Decision-Making and Reasoning Problem Performance..........59
Table 5: Mean Number of Features Generated in Consumer Task....................63
Table 6: Selection Task Performance Frequency Distribution in This Study and from Stanovich (1999; Stanovich & West, 1998)..................................................65
Table 7: Frequencies of Response Selections on the Charles in Scotland Problem..................................................................................................................67
Table 8: Mean Number of My-side and Other-side Arguments Generated and the Belief Bias Index in The Argument Generation Task.............................71
Table 9: Phi Coefficients for Formal Decision Making and Reasoning Tasks.....74
Table 10: Phi Coefficients with P and All Choosers in the Selection Task............79
Table 11: Domain Generality within Consumer Task: Correlations Between Number of Features Generated.................................................................82
Table 12: The Argument Generation Task: Correlations Between Total Number of Arguments Generated For Each Issue.............................................83
Table 13: Domain generality within Argument Generation Task: Correlations between Number of My-side and Other-side Arguments Generated..........................84
Table 14: Correlations Between Belief Bias Indices on The Argument Generation Task.................................................................86
Table 15: The Consumer Task and The Argument Generation Task: Correlations Between Features and Arguments Generated...............88
Table 16: Total Cognitive Ability, MFFT<sub>RT</sub>-Errors, and Need for Cognition Total Raw Score Means and Standard Deviations on Formal Tasks........................................................................92
Table 17: Summary of Cognitive Ability and Thinking Disposition Differences For All Formal Decision-Making and Reasoning Problems.................................................................93
Table 18: Cognitive Ability and Thinking Disposition Comparisons On The Prisoner’s Dilemma..............................................................94
Table 19: Cognitive Ability and Thinking Disposition Comparisons On Newcomb’s Problem.................................................................96
Table 20: Cognitive Ability and Thinking Disposition Comparisons On Disease Framing Problem.................................................................97
Table 21: Cognitive Ability and Thinking Disposition Comparisons on Box Problem........................................................................99
Table 22: Cognitive Ability and Thinking Disposition Comparisons On The Selection Task.................................................................102
Table 23: Cognitive Ability and Thinking Disposition Comparisons on Knights and Knaves Problem......................................................104
Table 24: Cognitive Ability and Thinking Disposition Differences Between Broadly Analytic (Cor+Partial+June/Kate) and Heuristic (Charles) responders On The Charles in Scotland Problem.................105
Table 25: Cognitive Ability and Thinking Disposition Differences Between Highly Analytic (Cor+Partial) and Undifferentiated Other
(June/Kate+Charles+Nothing+Blank+Other) Responders on the Charles in Scotland Problem

Table 26: Cognitive Ability and Thinking Disposition Differences Between Broadly Analytic (Cor+Partial+June/Kate) and Differentiated Other(Charles+Nothing+Blank+Other) Responders in The Charles in Scotland Problem

Table 27: Cognitive Ability and Thinking Disposition Differences on The Married Problem

Table 28: Cognitive Ability and Thinking Disposition Differences on The Green Levels Problem

Table 29: Simultaneous Regression Analyses With Formal Decision-Making and Reasoning Problems

Table 30: Correlations between Unique Features Generated and Cognitive Ability on Consumer Task

Table 31: Correlations between Unique Features Generated and Thinking Dispositions on Consumer Task

Table 32: Correlations between Consumer Reports Features Generated and Cognitive Ability on Consumer Task

Table 33: Correlations between Consumer Reports Features Generated and Thinking Dispositions on the Consumer Task

Table 34: Correlations between Argument Generation Task and Cognitive Ability

Table 35: Correlations between Argument Generation Task and Thinking Dispositions

Table 36: Belief Bias on The Argument Generation Task as a Function of Prior Opinion Strength and Knowledge
Table 37: Correlations Between Cognitive Ability and Thinking Dispositions Measures

Table 38: System 1 and System 2 Responses in Formal Tasks
List of Figures

Figure 1: Disease Framing Problem Decision Tree......................................................8
Figure 2: Decision Tree for Knights and Knaves Problem........................................10
Figure 3: Visual Schematic of Research Strategy and Relationships
of Interest.....................................................................................................................31
Figure 4: Frequency Distribution of the Number of Unique Features
Generated on The Consumer Task..............................................................................64
Figure 5: Frequency Distribution of Belief Bias Indices on
The Argument Generation Task...............................................................................73
Figure 6: Diagrammatic Conceptualization of the Domain Generality
Relations Between the Formal Tasks.....................................................................130
List of Appendices

Appendix A: Consumer Task.................................................................173
Appendix B: Scoring Guide for The Consumer Task..............................175
Appendix C: Argument Generation Task................................................182
Appendix D: Scoring of Charles in Scotland Formal Reasoning Problem.....185
Appendix E: Sample From the Matching Familiar Figures Test..................188
Appendix F: Scales Used In The Thinking Dispositions Questionnaire........189
Appendix G: Vocabulary Checklist..........................................................194
Appendix H: Order of Task Administration.............................................195
Introduction

The idea of considering the alternatives in a problem is a very common and pervasive way of describing an important part of good thinking (for example, see Nickerson, Perkins, & Smith, 1985; Bransford & Stein, 1993; Hayes, 1989; Beyer, 1988). For example, alternative consideration has been described as a key aspect in programs that claim to develop better thinking skills, such as Edward de Bono's Lateral Thinking. A lack of alternative consideration has also been purported to explain the causes of international conflict and war (Nicholson, 1997). Consideration of alternatives is clearly an essential skill that has been described in many domains, but this aspect of thinking has never been studied empirically in any systematic manner.

Shafir (1994) has suggested a term for the general concept of alternative consideration, and he has called it disjunctive thinking. Shafir terms disjunctive thinking as the tendency to consider all of the possible states of the world and to evaluate the consequence of each. Shafir and Tversky (1992) have also called a lack of alternative consideration as nonconsequentialism. Shafir (1994) likens disjunctive thinking to traveling through all of the branches of a decision tree. This type of thinking is fundamental if one wishes to maximize the utility of actions in decision-making as in, for example, expected utility theory. It is equally important in reasoning situations in which all possible intermediate problem states must be worked through in order to solve the problem. Nevertheless, research has shown surprising failures in people's performance on disjunctive problems (e.g., Shafir, 1994; Shafir & Tversky, 1992).

Shafir (1994) has discussed the concept of disjunctive thinking in the decision making and reasoning literatures, but these literatures are virtually independent. Although Shafir (1994) argues that disjunctive thinking is a
common skill across both the decision-making and reasoning domains, it is still unclear whether the concept reflects a common reasoning style across these two domains. One of the primary goals of this investigation was to examine the issue of domain generality. Specifically, would a common disjunctive thinking style generalize across the set of decision making and reasoning problems used in this study? The identification of a common thinking skill, such as disjunctive thinking, has many interesting implications. For example, it permits more common discourse across a broader set of disciplines, from educational thinking programs to examination of how decisions are made during international conflict.

The second goal of this study was to investigate who does well on disjunctive thinking problems from an individual differences perspective. This second purpose was based on Shafir’s (1994) surprising report of poor performance on these problems. The framework for the individual differences investigation was derived from Baron’s (1985) distinction between cognitive capacities and thinking dispositions and Stanovich’s (1999; Stanovich & West, 2000) Generic Dual Process Framework. This distinction between cognitive capacities and thinking dispositions is crucial, as both Baron (1985) and Stanovich (1999) view thinking dispositions as more malleable than cognitive capacities. It was of potential educational importance to examine whether individual differences in disjunctive reasoning and decision-making would be associated with thinking dispositions after the variance explained by cognitive capacity had been partialled out.

The Domain Generality of Disjunctive Thinking: A Taxonomy of Tasks

Disjunctive thinking has been studied in the reasoning literature and in the decision making literature—but rarely together (see Shafir, 1994, for an exception). In fact, several investigators have commented upon the relatively
isolated and separate development of the decision theory and reasoning literatures (e.g., Evans, Over, & Manktelow, 1993). For example, Evans, Newstead, and Byrne's (1993) analysis of disjunctive reasoning is limited to the study of deductive logic and the construction of truth tables. Also, complete texts on the Prisoner's Dilemma, a decision-making problem, have been devoted to the study of this single problem (for example, see Campbell & Sowden, 1985), but virtually no attention has been given to the relationship between performance on the Prisoner's Dilemma and other reasoning problems. Shafir and Tversky (1992) emphasize disjunctive thinking in decision making situations and in reasoning problems. Some studies have begun to make attempts to merge the reasoning and decision-making literatures. For example, Legrenzi, Girotto, and Johnson-Laird (1993) used a mental models framework to examine performance on deductive syllogisms, a Wason card selection task, and in a decision-making task. In this study, an even broader set of formal reasoning and decision-making problems was employed. In addition, some informal tasks were also included in order to test the generality of a disjunctive thinking skill on tasks that model more real world decision-making and reasoning. Shafir's (1994) discussion of decision under uncertainty and thinking through disjunctions formed the basis for the taxonomy of tasks used in this study. The tasks that Shafir (1994) included in his discussion are presented in Table 1.

**Disjunctive Consideration in Decision-Making Problems**

A very compelling demonstration of the failure to think disjunctively was provided in a problem studied by Tversky and Shafir (1992). Termed the Hawaii problem, participants were told to imagine that they have just taken a tough examination at the end of a semester. They were presented with one of two scenarios. In one scenario they found out that they had passed the exam.
Shafir (1994) identified the following problems as instances of disjunctive thinking:

**Decisions:**
- Risky Choice Problem
- The Hawaii Problem

**Games:**
- Newcomb
- Prisoner's Dilemma

**Probabilistic Judgment:**
- Box Problem

**Puzzles and Paradoxes:**
- The Impossible Barber
- Knights and Knaves

**Deductive Inference:**
- Wason Selection Task
- THOG Problem
- Charles in Scotland Problem

In the other scenario they found out that they had failed the exam. Then participants were told that they had an opportunity to purchase a very attractive Christmas vacation package at an exceptionally low price. The question was whether they would decide to buy the vacation package. Tversky and Shafir (1992) reported that 54% of their participants indicated that they would buy the package if they passed the exam, and 57% indicated that they would buy the package if they failed the exam. Therefore, in either the pass or fail circumstance, over half of participants indicated that they would buy the vacation package.

In another version the Hawaii Problem, another group of participants were given the same problem, except that they had to make a decision without knowing whether they passed or failed their exam. The results were astoundingly different. In this version, only 32% of participants indicated that they would buy the package. It was additionally found that 61% indicated that they would pay a $5 non-refundable fee in order to defer the decision until they
found out whether they passed or failed the exam. Interestingly, a majority of participants reported that they would buy if they passed and a majority would buy if they failed. When participants did not know the outcome of their examination, a majority chose not to buy. Significantly fewer participants opted to buy the vacation package when the outcome of their examination was uncertain. How can we explain these seeming discrepant results?

Shafir and Tversky (1992) call this phenomenon the disjunction effect, and state that a disjunction effect occurs when “people prefer X over Y when they know that A obtains, and they also prefer X over Y when that event A does not obtain, but they prefer Y over X when it is unknown whether or not A obtains” (p. 451). The disjunction effect is a violation of one of the most basic principles of rational choice, called Savage’s Sure Thing Principle (STP). The Sure Thing Principle states:

If a person would prefer a to b knowing that X obtained, and if he would prefer a to b knowing that X did not obtain, then he definitely prefers a to b (Shafir, 1994, p. 404).

Shafir and Tversky (1992) attribute violations of the Sure Thing Principle to situations in which uncertainty is present. They state that “uncertain situations may be thought of as disjunctions of possible states: either one state will obtain, or another” (Shafir & Tversky; 1992; p. 449-450). These types of problems can be likened to a hypothetical decision tree, and that people need to exhaustively travel all of the branches of this tree to make the best decision. If individuals actually worked through all of the disjuncts, they would realize their inconsistency in preference, and consequently should not show the disjunctive effect. The disjunction effect on the Hawaii Problem has been demonstrated in within subject designs as well as between subject designs (Tversky & Shafir, 1992).
Disjunctive Thinking

Framing effects also come about because of the failure to think disjunctively. Framing problems are another instance in which the disjunction effect is displayed. Framing problems have been studied in the decision-making literature (Tversky & Kahneman, 1981; Kahneman & Tversky, 1982), and they have been researched more extensively than the Hawaii Problem. Tversky and Kahneman (1981) have demonstrated framing effects, for example, in the Disease Framing Problem. This problem is as follows:

**Problem 1: 'Lives saved' Version.** Imagine that the United States is preparing for the outbreak of an unusual Asian disease expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

a) If Program A is adopted, 200 people will be saved.

b) If Program B is adopted, there is a 1/3 probability that 600 people will be saved, and a 2/3 probability that no people will be saved.

If it was your decision to choose between Program A or Program B, which program would you choose?

**Problem 2: 'Lives Lost' Version.** Imagine that the United States is preparing for the outbreak of an unusual Asian disease expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

a) If Program C is adopted, 400 people will die.

b) If Program D is adopted, there is a 1/3 probability that nobody will die, and a 2/3 probability that 600 people will die.

If it was your decision to choose between Program C or Program D, which program would you choose?

In the Disease Problem, participants are presented with two different versions of the problem. Participants are asked to select Program A or B in Problem 1, and to select Program C or Program D in Problem 2. Figure 1 displays a decision tree to illustrate the disjunctive thinking involved in this problem. As shown in Figure 1, disjunctive thinking would involve constructing a single representation of both versions of the problem. That is, integrating the Lives
Disjunctive Thinking

Saved version and the Lives Lost version into one representation. There are two possible disjunctive solutions. One disjunctive solution would be the selection of Program A in the first version and Program C in the second version, and this is called the risk averse solution. That is, selection of Program A and Program C both reflect an aversity to risk, as the outcomes are stated in absolute values as opposed to probabilities in both options. The other disjunctive solution would be the selection of Program B in the first version and Program D in the second version, and this is called the risk-taking solution. Counter to the risk averse solution, selections of Program B and Program D reflect risk-seeking, as both outcomes are stated in probabilities, as opposed to absolute values.

Tversky and Kahneman (1981) reported 73% of respondents selected Program A in Problem 1 and Program D in Problem 2, which is a non-disjunctive solution. This has been called the framing effect in the literature because the selection of Program A reflects a response tendency to be risk averse, and the selection of Program D reflects a tendency to be risk taking. Kahneman and Tversky (1982) use prospect theory to explain why framing effects occur. According to prospect theory (Kahneman & Tversky, 1982), the framing effect occurs because of the S-shaped value function that characterizes how respondents evaluate risky prospects. In particular, the response to losses is much more extreme than the response to gains. Tversky and Kahneman (1981) state that: "the displeasure associated with losing a sum of money is generally greater than the pleasure associated with winning the same amount" (p. 454). According to prospect theory, a rational decision-maker prefers the prospect that offers the highest expected utility (Fischhoff, 1983). Prospect theory is a descriptive theory that explains how and why choices deviate from the normative model (Baron, 1994). In the case of the
Disjunctive Thinking

Disease Problem, participants attribute different utilities depending on whether outcomes are described as gains or losses (Baron, 1994). The incongruity in program choices on the Disease Problem is akin to the disjunction effect, as described by Shafir (1994). In the Disease Problem, exhaustively traveling each branch of the decision tree should lead one to realize that the expected values are equivalent for each option. Figure 1 displays the disjunctive options in the Disease Problem in the form of a decision tree.

**Figure 1. Disease Framing Problem Decision Tree**

As shown in Figure 1, the realization that both options lead to equivalent outcomes should lead the rational decision-maker to be indifferent about the two programs. A recognition of this equivalence is captured in this problem by the consistent selection of the options reflecting the tendency to be risk averse.
or the tendency to be risk-taking. An inconsistent selection, that is, risk aversity in one version and risk-taking in the second version, is therefore an indication of a lack of having traveled all of the branches of the decision tree.

Other decision making problems that were used in this study were the Box Problem, the Prisoner's Dilemma, and Newcomb's Problem, all from Shafir (1994).

**Disjunctive Consideration in Reasoning Problems**

Disjunctive reasoning has mainly been described in the deductive reasoning literature (for example, see Evans, Newstead, & Byrne, 1993). The focus of the disjunctive reasoning literature has been on people's understanding of the logic of disjuncts, for example, whether they understand inclusive and exclusive disjuncts (Evans, Newstead, & Byrne, 1993). This literature has not treated disjunctive thinking as the expected utility or outcome of events, as in decision theory (Evans, Over, & Manktelow, 1993).

Some of the typical problems studied in the reasoning literature have been the Selection Task, the Knights and Knaves Problem, and the Charles in Scotland Problem. All of these problems are characterized by disjunctive consideration, but even these problems have never been studied together empirically. Variants of the Selection Task, for example, have a long history of systematic study (Evans, Newstead, & Byrne, 1993), but it is unclear whether the phenomenon captured by this task is unique to this problem or whether the pattern of results are shared with other reasoning problems. The Knights and Knaves Problem has been studied less extensively than the Selection Task, but it has been studied in some detail by Rips (1989), and also by Shafir (1994). The prototypical Knights and Knaves Problem that Shafir (1994) and Rips (1989) discuss is as follows:
Knights and Knaves
There are 3 inhabitants, A, B, and C, each of whom is a knight or a knave. Knights always tell the truth, while knaves always lie. Two people are said to be of the same type if they are both knights or knaves. A and B make the following statements:

A: B is a knave
B: A and C are of the same type
What is C? (from Shafir, 1994, p. 424)

A decision tree illustrating the disjunctive nature of this problem is displayed in Figure 2.

Figure 2. Decision Tree for Knights and Knaves Problem

As displayed in Figure 2, if one travels down both branches of this tree, one realizes that there is only one possible answer: person C can only be a knave.
Disjunctive Thinking

Only about 20 to 32% of participants solve the problem correctly (Rips, 1989). Optimal performance on the Knights and Knaves Problem is very low, and it is of interest to determine why performance is so low, and where participants go wrong in this problem. Rips (1989) presented participants with many versions of the Knights and Knaves problems ranging in complexity and the number of steps involved in solving the problem. Rips (1989) reported that as the number of assumptions or inference steps increased in each problem, it took longer for participants to solve the Knights and Knaves problems, and fewer participants actually solved the more difficult versions correctly. None of participants solved the most difficult version of the problem, and 35% of participants solved the easiest version of the problem (Rips, 1989). Rips' (1989) easiest version of the Knights and Knaves Problem was used in this study. Rips (1989) used a natural deduction model to explain performance on this problem. The natural deduction model is based on the idea that participants use a systematic set of deductive rules to solve the problem. Rips (1989) noted one limitation of the natural deduction model to explain performance, and that was the huge variation in performance across participants. Rips (1989) found that performance of individual participants varied between 0% to 84% correct. Rips suggested that either his explanation only applies to a select subset of participants who were successful at solving the problems, or that a motivational explanation might account for these differences. Rips (1989) did not elaborate or report any data relevant to his speculations about individual differences.

Johnson-Laird and Byrne (1990) proposed the mental models theory as a better alternative to account for Rips' (1989) findings with the Knights and Knaves Problem. Participants in Rips' (1989) study performed worse when the number of steps were increased and when negated conjunctions were
Disjunctive Thinking

included in the problem. Johnson-Laird and Byrne (1990) address these findings by stating that:

ordinary individuals have a limited ability to process models of premises. Hence, they cannot cope with negated conjunctions that force them to consider the consequences of a disjunctive set of models, they have only a limited ability to follow up the consequences of assumptions, and they find positive matches easier than negative mismatches. (p. 77)

Johnson-Laird's mental models theory (1983) posits that performance on problems that require the representation of many models is constrained by the limited capacity of working memory. The limited working memory capacity explanation has been used as a frequent explanation of performance on other reasoning problems, such as performance on deductive syllogisms and the Wason Selection Task (Legrenzi, Girotto, & Johnson-Laird, 1993), and on action and inference problems (Mynatt, Doherty, & Dragan, 1993). Shafir (1994) has also suggested that poor performance on the Knights and Knaves Problem may be attributed to the complexity of the problem. Rips suggested that poor performance may be attributed to difficulty with "conceptual bookkeeping" as opposed to a "logical deficiency" (Rips, 1989). Similar explanations alluding to limited ability have been used to explain performance on other classic reasoning problems, for example, on the THOG problem (Evans, Newstead, & Byrne, 1993; Wason & Brooks, 1979).

The Knights and Knaves Problem illustrates the concept of disjunctive thinking, as described by Shafir (1994). Other similar problems that were used in this study included the Charles in Scotland Problem (Shafir, 1994; Baron, 1995), an abstract or non-deontic version of the Selection Task (Wason, 1966; Evans & Over, 1996; Newstead & Evans, 1995; Evans, Newstead, & Byrne, 1993; Stanovich & West, 1998; Stanovich, 1999), the Married Problem (Levesque, 1986), and the Green Levels Problem (Levesque, 1989).
Disjunctive Thinking

Logic of the Disjunctive Tasks Used in this Study

The present study examined two broad domains of tasks in which disjunctive thinking has been implicated—the decision making literature and the reasoning literature. Within each of these categories of tasks, several different formal measures and more realistic informal reasoning measures were employed. The extent of domain generality across each of the formal and informal measures within each of the categories was assessed. Also, the extent of domain generality across the broad decision making and reasoning categories was assessed. Table 2 lists the formal reasoning and decision-making tasks that were used in this study. Descriptions of each task are discussed in detail in the Method section.

Table 2

Formal and Informal Reasoning and Decision-Making Tasks Used in This Study

<table>
<thead>
<tr>
<th>Formal Decision-Making Problems</th>
<th>Formal Reasoning Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prisoner's Dilemma</td>
<td>Selection Task</td>
</tr>
<tr>
<td>Newcomb's Problem</td>
<td>Knights and Knaves</td>
</tr>
<tr>
<td>Disease Framing Problem</td>
<td>Charles in Scotland</td>
</tr>
<tr>
<td>Box Problem</td>
<td>Married Problem</td>
</tr>
<tr>
<td></td>
<td>Green Levels Problem</td>
</tr>
<tr>
<td>Informal Decision-Making Task</td>
<td>Informal Reasoning Task</td>
</tr>
<tr>
<td>The Consumer Task</td>
<td>The Argument Generation Task</td>
</tr>
</tbody>
</table>

The second purpose of this study was to examine who responds disjunctively on the decision-making and reasoning tasks used in this study. An individual differences perspective was used to address this purpose. In particular, the framework for the individual differences investigation was derived from Baron's (1985) distinction between cognitive capacities and
disjunctive thinking and Stanovich's (1999; Stanovich & West, 2000) Generic Dual Process Framework. If response tendencies on the individual difference measures are related on some or all of the disjunctive problems, such a finding would serve both purposes of the study. First, systematic differences in the individual difference measures on the decision-making and reasoning problems would provide some explanations as to who does well on the types of problems used in this study. In particular, systematic differences on the thinking dispositions would implicate correlates which are more susceptible to change and remediation, as compared to cognitive capacities which are less susceptible to remediation. Second, trends or patterns of individual differences across some or all of the decision-making and reasoning tasks would provide some converging evidence for the first purpose of this study. That is, patterns of individual differences across the different tasks would provide converging evidence that these problems capture a similar thinking skill.

What Factors are Associated with Individual Differences in Disjunctive Thinking? The Distinction Between Cognitive Capacities and Thinking Dispositions

One central issue in the literature on critical thinking is whether individual differences in reasoning performance are associated with cognitive capacities or thinking dispositions (Baron, 1991, 1994; Perkins, Jay, & Tishman, 1993). A major focus in the literature has been on the separability of cognitive capacities and thinking dispositions, and many have discussed the overlap between these individual differences (for example, Baron, 1985). Baron (1994) argues for the separability of thinking dispositions and cognitive capacities. Stanovich (1999) argues that cognitive capacities and thinking dispositions are constructs which can be placed at different levels of analysis in cognitive theory. The separability of cognitive ability and thinking
dispositions was of interest in the current investigation.

Cognitive capacities are an individual difference measure which include abilities such as processing speed, speed of retrieval, and working memory capacity (Baron, 1994). Cognitive capacities tend to be less malleable in an individual, and Baron (1994) argues that capacities are likely to be determined by genetic and physiological factors rather than by learning. It is often argued that these types of abilities constitute and explain a large component of intelligence (Baron, 1985; 1994). Limits in cognitive capacities or computational limitations are often used to explain reasoning performance, and moderate differences are typically found to explain reasoning performance (Stanovich, 1999; Stanovich & West, 2000; Stanovich & West, 1998; Johnson-Laird & Byrne, 1990; Mynatt, Doherty, & Dragan, 1993; Legrenzi, Girotto, & Johnson-Laird, 1993; Evans, Newstead, & Byrne, 1993). The reliance on cognitive capacities as an explanation of reasoning performance is controversial (Baron, 1991), and it has been suggested that the definition of cognitive abilities or intelligence should be expanded to include more modifiable and controllable abilities (Haslam & Baron, 1994).

In contrast, thinking dispositions have been defined as "learned tendencies to behave in certain ways" (Baron, 1991, p. 118). Thinking dispositions are not fixed and can be somewhat altered with practice and across different situations (Baron, 1991). Thinking dispositions have also been viewed as cognitive styles that are more malleable and teachable (Stanovich, 1999). For example, it has been shown that training in the cognitive style of reflectivity improved children's performance on school tasks (Baron, Badgio, & Gaskins, 1985), and adolescents' performance on a conditional reasoning task (Overton, Byrnes, & O'Brien, 1985). Many different cognitive styles pertaining to reasoning performance have been discussed in the literature, for
example, the reflectivity-impulsivity style developed by Kagan, Rosman, Day, Albert, and Phillips (1964), the need for cognition style (Cacioppo, Petty, Feinstein, & Jarvis, 1996), and the actively open-minded thinking style (Stanovich & West, 1997).

It has been argued that a strict cognitive capacity approach is a very limited way of explaining reasoning performance (Haslam & Baron, 1994), as cognitive capacities cannot generally be improved by practice. It has been demonstrated that cognitive capacities are correlated with thinking and reasoning abilities, and this same evidence has also demonstrated that there is variance left over that is not accounted for when cognitive ability or IQ is partialled out (for example, Sá, West, & Stanovich, 1999; Stanovich, 1999; Stanovich & West, 1997). Such evidence would lead one to suggest that IQ is not the only factor involved.

Perkins, Farady, and Bushey (1991) also report evidence which demonstrates that IQ does not entirely account for good reasoning. In this study, participants were presented with a social issue question, such as: "Would providing more money for public schools significantly improve the quality of teaching and learning?" (Perkins et al., 1991). Then, participants were asked to give their perspectives on the issue, and to report their reasons for why they supported a given side. In short, Perkins et al. (1991) found that high IQ was significantly correlated with the number of my-side arguments generated, but no such effect was found between IQ and other-side arguments. Thus, it seems that IQ contributed to reasoning ability, but these effects also suggest that IQ might be used to select information that supports a preferred position. Perkins et al. (1991) suggest that "having a high IQ gives no guarantee of one's using it well anymore than having a lot of horsepower under the hood of your car guarantees your driving the car well" (p. 103).
The distinction between capacities and dispositions has also been characterized as a difference between the algorithmic and rational levels of explanation (Stanovich & West, 1997), which has been described in the Generic Dual Process Framework for reasoning proposed by Stanovich (1999; Stanovich & West, 2000). Different levels of analysis have been used in cognitive theory in order to explain human behaviours and functions, and this approach has been proposed by many theorists, such as Anderson (1990), Marr (1982), Newell (1982, 1990), and Dennett (1978, 1987). As discussed by Stanovich and West (1997), Anderson (1990) defines four main levels of analysis: the biological level, the implementation level, the algorithmic level, and the rational level. The biological and implementation levels pertain to the hardware and specifications of the hardware level that are not as relevant to cognitive theory. The algorithmic level pertains to specifying the computational processes needed to carry out a task (Stanovich & West, 2000; Stanovich, 1999). The rational level refers to specifying the goals of the system, namely what the system is doing and why (Stanovich & West, 2000; Stanovich, 1999). Stanovich and West (1997) argued that individual differences in cognitive capacities can be attributed to the functions at the algorithmic level of explanation, while variations in thinking dispositions can be attributed to a rational level of explanation. Stanovich and West (1997) found that thinking dispositions predicted further variation on performance on an informal reasoning task after cognitive capacities were partialled out. They explained this additional contribution of thinking dispositions by suggesting that good reasoning is not only explicable by cognitive capacities at the algorithmic level of explanation, but that the rational level is also important, as “we must know something about the epistemic goals of the reasoners” to understand reasoning performance (Stanovich & West, 1997, p. 352).
Originating from the levels of analysis perspective, Stanovich (1999) used a generic Generic Dual Process Framework to integrate the roles of cognitive ability and thinking dispositions in explaining reasoning performance. This framework captures many aspects of previous dual process models, and Table 5.1 from Stanovich (1999; p. 145) is replicated below in Table 3 in order to facilitate the explanation of this framework. System 1 and System 2 in this framework are generic terms used to explain two very different modes of thinking. The parameters of this Generic Dual Process Framework integrate or implicate many of the classic theories of reasoning, including the mental models approach (Johnson-Laird, 1985, 1994, 1999; Johnson-Laird, Byrne, & Schaeken, 1994), rule-based approaches (Rips, 1989; O'Brien, 1995), heuristic-analytic models (Evans, 1984, 1989), and pragmatic reasoning schemas (Holyoak & Cheng, 1995; Cheng & Holyoak, 1985). System 1 is viewed as encompassing interactional forms of intelligence and is associated with highly contextualized, relatively quick, and heuristic processing. On the other hand, System 2 encompasses analytic forms of intelligence which are decontextualized, rule-based, more controlled, and more demanding of cognitive capacities. System 2 types of processing have most often been linked with traditional measures of psychometric intelligence. The advantage of System 1 is speed, and the advantage of System 2 is flexibility. Both of these systems are important for different reasons, and dysfunctions in either System 1 (Damasio, 1994) or System 2 (Stanovich, 1999) can lead to non-adaptive or non-optimal outcomes. System 1 was be viewed as the default system, or heuristic, that is cued automatically, but System 2 is not automatically cued.

The placement of analytic processes in System 2 functions has important implications for predictions regarding cognitive ability differences.
Table 3

The Terms for the Two Systems Used by a Variety of Theorists and the Properties of Dual-Process Theories of Reasoning (reproduced from Stanovich, 1999)

<table>
<thead>
<tr>
<th>Dual-Process Theories:</th>
<th>System 1</th>
<th>System 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sloman (1996)</td>
<td>associative system</td>
<td>rule-based system</td>
</tr>
<tr>
<td>Evans (1984, 1989)</td>
<td>heuristic processing</td>
<td>analytic processing</td>
</tr>
<tr>
<td>Evans &amp; Over (1996)</td>
<td>tacit thought processes</td>
<td>explicit thought</td>
</tr>
<tr>
<td>Reber (1993)</td>
<td>implicit cognition</td>
<td>explicit learning</td>
</tr>
<tr>
<td>Levinson (1995)</td>
<td>interactional intelligence</td>
<td>analytic intelligence</td>
</tr>
<tr>
<td>Epstein (1994)</td>
<td>experiential system</td>
<td>rational system</td>
</tr>
<tr>
<td>Pollock (1991)</td>
<td>quick &amp; inflexible modules</td>
<td>intellecution</td>
</tr>
<tr>
<td>Hammond (1996)</td>
<td>intuitive cognition</td>
<td>analytical cognition</td>
</tr>
<tr>
<td>Klein (1998)</td>
<td>recognition-primed decisions</td>
<td>rational choice strategy</td>
</tr>
<tr>
<td>Properties:</td>
<td>associative</td>
<td>rule-based</td>
</tr>
<tr>
<td></td>
<td>holistic</td>
<td>analytic</td>
</tr>
<tr>
<td></td>
<td>automatic</td>
<td>controlled</td>
</tr>
<tr>
<td></td>
<td>relatively undemanding of cognitive capacity</td>
<td>relatively slow</td>
</tr>
<tr>
<td></td>
<td>relatively fast</td>
<td>acquisition by cultural and formal tuition</td>
</tr>
<tr>
<td></td>
<td>acquisition by biology, exposure,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>and personal experience</td>
<td></td>
</tr>
<tr>
<td>Task Construal:</td>
<td>highly contextualized</td>
<td>decontextualized</td>
</tr>
<tr>
<td></td>
<td>personalized</td>
<td>depersonalized</td>
</tr>
<tr>
<td></td>
<td>conversational and socialized</td>
<td>asocial</td>
</tr>
<tr>
<td>Type of Intelligence Indexed:</td>
<td>interactional (conversational implicature)</td>
<td>analytic (psychometric IQ)</td>
</tr>
</tbody>
</table>

In particular, it would be expected that those who have higher levels of cognitive ability would be more likely to display optimal performance on reasoning tasks that involve System 2 type processes. Stanovich (1999) and Stanovich and West (2000) report strong evidence that this in fact the case on a
number of reasoning tasks, including performance on syllogisms, the
selection task, a statistical reasoning task, and an argument evaluation task.
These findings are consistent with discussions of some of the disjunctive tasks
that were used in this study. For example, cognitive ability differences
(Johnson-Laird & Byrne, 1990) and rule-based processes associated with
System 2 (Rips, 1989) have been implicated in performance on the Knights and
Knaves Problem (see previous discussion of Knights and Knaves Problem).
Cognitive ability, IQ, or analytic intelligence has been studied extensively in
the reasoning literature, and therefore positive manifold may be expected
(Stanovich, 1999). That is, cognitive ability will likely be correlated with
performance on the reasoning tasks used in the present study.

In addition to explorations of cognitive ability, the present investigation
built on previous theoretical suggestions that there is systematic variance left
over (Stanovich, 1999). Stanovich (1999; Stanovich & West, 1998), for example,
reported evidence that thinking dispositions, such as the actively open-minded
thinking (AOT) scale, predicted additional variance after cognitive ability had
been partialled out. Additional variance was predicted on a series of tasks,
including an informal reasoning task, syllogisms, statistical reasoning,
covariation detection, hypothesis testing bias, and outcome bias tasks
(Stanovich & West, 1998b). Stanovich (1999; Stanovich & West, 1997) suggested
that the thinking dispositions that were used tapped tendencies toward
epistemic self-regulation. For example, the actively open-minded thinking
scale (AOT) involves tendencies to consider evidence that goes against one's
beliefs, and a willingness to consider alternative perspectives (Baron, 1985).
The characterization of System 2 processes as slow, controlled, and
decontextualized captures tendencies of thinking dispositions, such as the
AOT. Others have also reported a unique contribution of thinking styles or
Disjunctive Thinking

Dispositions after cognitive ability was controlled. For example, Cacioppo, Petty, Kao, and Rodriguez (1986) found that the need for cognition disposition and verbal intelligence significantly predicted the number of arguments that participants could recall after being presented with a persuasive message. After verbal intelligence was statistically controlled, the need for cognition style accounted for significant additional variance in the number of arguments recalled.

In summary, all of these studies demonstrate that cognitive capacities are related to reasoning performance. These studies also demonstrate that there is unexplained variance that is not accounted for by cognitive capacities or analytic intelligence. Some evidence has shown that thinking dispositions predict performance on many reasoning tasks, including argument evaluation (Stanovich & West, 1997), and argument recall (Cacioppo, Petty, Kao, & Rodriguez, 1986).

The role of thinking dispositions in explaining performance on some disjunctive thinking types of problems has been considered. Smith & Levin (1996), for example, examined the need for cognition disposition on framing problem performance. In a Life in Danger version of a framing problem, Smith and Levin found that individuals who had a higher score on the need for cognition scale were less likely to display a framing effect than those who had a lower score on the need for cognition scale. Smith and Levin’s (1996) discussion is consistent with our conceptualization of dispositional tendencies in this study, as they state that: “the more thought one devotes to a decision, the less likely the distortion of that choice by framing effects” (p. 288).

Consistent with this finding is the notion of reflectivity, or spending more time with a problem. Takemura (1992), for example, found that participants who were asked to spend more time on an Investment Framing problem were less
likely to display a framing effect.

Jou, Shanteau, and Harris (1996) also discussed the role of individual differences in framing problems, more specifically, cognitive ability limitations:

"Many of the computations may well be beyond the capacity of ordinary people's mental calculation. People may selectively attend to some salient aspects of such complex numerical information and overlook others because a complete computation of the outcomes exceeds their attentional and short-term memory span. Thus, some judgmental inconsistencies under those circumstances may derive from performance limitations rather than from some fundamental deficiencies in the competence of the human cognitive system." (p. 10).

In their study, Jou et al. (1996) found that significantly fewer participants displayed a framing effect when they were presented with rationales, or causal schemas, in the framing problems. They argued that causal schemas helped to bring responders' attention to different aspects of the problem, aspects that may be more implicit and less obvious. Causal schemas may have served to further engage responders, which is one of the functions associated with thinking dispositions. Further, Jou et al. (1996) also imply a role for thinking dispositions based on their discussion of causal schemas:

From the information-processing point of view, framing is a form of manipulating the salience or accessibility of different aspects of information. One aspect of a consequence can be highlighted and the other de-emphasized, or even equivocated. A decision is the product of the combined impressions about various aspects of information which are often unequally weighted (Birnbaum & Jou, 1990). The function of introducing the rationale is to ensure that the underweighted side of the information if given greater weight. (p. 9)

If elaboration of a schema is a function of considering alternatives, then it would be expected that thinking dispositions, such as reflectivity, should play a role in performance on framing problems.

In summary, both cognitive capacities and thinking dispositions have been implicated in explanations of performance on reasoning tasks. As the
focus in the cognitive literature has been on cognitive capacities, that is, the algorithmic level of explanation, one main purpose of this study was to examine the role of thinking dispositions in disjunctive thinking. A series of thinking dispositions that have previously been used in the literature were selected based on how well they matched or described the processes involved in disjunctive reasoning, namely, the processes of generation and consideration of all of the alternatives.

**Thinking Dispositions Examined in the Present Investigation**

The thinking dispositions chosen for investigation in the present study were those thought to have particularly strong relationships with disjunctive thinking, as outlined above. Three main sets of thinking dispositions were investigated, including, the style of impulsivity and reflectivity (Kagan, Rosman, Day, Albert, & Phillips, 1964), the need for cognition style (Cacioppo, Petty, Feinstein, & Jarvis, 1996), and actively open-minded thinking scale (Stanovich & West, 1998).

**The Dimension of Reflectivity and Impulsivity.** The dimension of reflectivity and impulsivity has a long history in the literature as a cognitive style (Messer, 1976). The classic task used to measure this dimension is the Matching Familiar Figures Test (MFFT), which was developed by Kagan, Rosman, Day, Albert, and Phillips (1964). In this task, participants are asked to match a picture target with one of four, six, or eight other picture stimuli. This task effectively captures a trade-off between speed and accuracy. That is, those who perform quickly and make many errors are labeled as impulsive, while those who respond more slowly and make fewer errors are labeled as reflective (Baron, Badgio, & Gaskins, 1986). The reflectivity-impulsivity dimension has mainly been studied in children (Baron, 1995), but it has also been used in studies with adults (Davidson, 1984, 1988; Van den Broek,
Bradshaw, & Szabadi, 1987; Denney & List, 1979; Van Merriëenkoer, Jelsma, Timmermans, & Sikken, 1990).

There has been some debate about whether the MFFT captures something stylistic and dispositional or whether performance is confounded by cognitive ability. The dimension of reflectivity and impulsivity has been described as a disposition, as people have some control over how quickly they want to get through a task (Baron et al., 1986). It has also been suggested that to be reflective is a preference for an analytic strategy (Baron et al., 1986). Kemler Nelson and Smith (1989) suggested that impulsivity is related to more holistic processing, while reflectivity is related to analytic thinking. However, van den Broek et al. (1987) found that mean latency to the first response was not significantly correlated with full-scale intelligence on the WAIS-R short form administration. This debate has a long history, particularly in the developmental psychology research literature (Egeland & Weinberg, 1976; Block, Block, & Harrington, 1974).

The capacity versus style debate is likely responsible for the decreased popularity and use of this task. More recent work has replicated the significant correlation between errors on the MFFT and intelligence, at least in children (Welsh, Pennington, & Groisser, 1991). Schachar and Logan (1990) also report that performance on MFFT depends on factors other than cognitive impulsivity, such as, intelligence, search strategy, metacognitive awareness of inhibiting the response until an exhaustive search has been completed, and decision criteria, motivation, or attentional capacity. Some of the factors listed by Schachar and Logan are related to capacity limitations, such as attentional capacity and intelligence, while others may be related to stylistic or dispositional factors, such as search strategy and decision criteria. Clearly, both ability and dispositional explanations have been used to describe
performance on the MFFT, and the critical issue is the separability of these
two constructs. Salkind and Wright (1977) have proposed statistical methods to
separate the ability and stylistic features of the MFFT. Salkind and Wright
(1977) provide statistical alternatives which increase the legitimacy of using
the MFFT as a style construct. Importantly, the analyses in the present
investigation took into account the confounding role of cognitive ability.

The MFFT captures the notion of reflectivity (Baron, 1985), a tendency to
reflect when several possible alternatives are available (Messer, 1976), and
attention to detail (Overton, Byrnes, & O'Brien, 1985). The reaction time
measure captures the tendency to spend more time on a task. The total
number of errors measure captures a form of response inhibition or
regulation which would seem to tap into knowing when to give a response. In
a review by Pennington and Ozonoff (1996), they reported that the MFFT has
been used as a measure of executive function, and that both reaction time and
errors have been demonstrated to be statistically sensitive. The selection of this
disposition was based mainly on the theoretical ideas underlying reflectivity
and impulsivity. Namely, the idea that careful and thoughtful consideration of
the alternatives would lead to more optimal performance on the tasks
involving disjunctive thinking.

**The Need For Cognition.** Another relevant thinking disposition, or
cognitive style, is called the need for cognition (NC) style. The “need for
cognition” is defined as follows:

Individuals high in need for cognition were proposed to naturally
tend to seek, acquire, think about, and reflect back on information
to make sense of stimuli, relationships, and events in their world.
Individuals low in need for cognition, in contrast, were characterized
as more likely to rely on others (e.g., celebrities and experts), cognitive
heuristics, or social comparison processes to provide this structure.
(Cacioppo, Petty, Feinstein, & Jarvis, 1996, p. 198)
The need for cognition dimension is measured through a self-report scale, and participants are asked to respond to whether a series of statements are true of them, such as: "I would prefer complex to simple problems" (Cacioppo et al., 1996). A high need for cognition has been correlated with other individual difference measures, such as cognitive innovativeness and a desire for new experiences that stimulate thinking (Venkatraman, Marlino, Kardes, & Sklar, 1990; Venkatraman & Price, 1990), intrinsic motivation (Olson, Camp, & Fuller, 1984; Amabile, Hill, Hennessey, & Tighe, 1994), a need to evaluate (Jarvis & Petty, 1996), and openness to experience (Berzonsky & Sullivan, 1992).

Petty and Cacioppo (1986) examined the need for cognition style in an informal reasoning task. They examined whether those who have a high need for cognition would generate thoughts that better reflect the quality of an argument presented than those who have a low need for cognition. Petty and Cacioppo (1986) examined this question by presenting college participants with a message about raising tuition. What they found is that the profile of thoughts generated by individuals high in need for cognition were more likely to reflect the quality of the arguments in the message than the profile of thoughts generated by individuals low in need for cognition (Cacioppo et al., 1996). That is, those who were high in need for cognition generated more positive evaluations when a strong argument was presented, and more negative evaluations when a weak argument was presented. Those who were low in need for cognition did not show this effect, and the thoughts they generated reflected their favourable or unfavourable attitudes toward the issue or argument. This study demonstrated that those who are high in need for cognition were sensitive to the quality of the argument, and framed their thoughts in the context of the argument presented. Thus, it seems that the
need for cognition style captures reflectivity and engagement with an argument, and a lower susceptibility to being captured by superficial cues, which is a form of decontextualization (Stanovich, 1999). These tendencies or dispositions are also important in disjunctive thinking. A similar result with the need for cognition disposition was reported on another informal reasoning task by Cacioppo, Petty, Kao, and Rodriguez (1986).

A disposition that is highly related to the need for cognition is the style of typical intellectual engagement (Ackerman & Heggestad, 1997). The style of typical intellectual engagement has been studied extensively in the personality literature (Ackerman & Heggestad, 1997). This self-report scale is composed of items such as: “I enjoy the challenge of reading a complicated novel,” and “I always feel that I must look into all sides of a problem.” The typical intellectual engagement scale conceptually captures the same tendencies toward reflectivity and engagement as in the need for cognition scale. The need for cognition scale and the typical intellectual engagement scale overlap considerably, and therefore they were collapsed in this study.

The development of the typical intellectual engagement scale is theoretically significant because it contributes to the distinction between cognitive capacities and thinking dispositions. Ackerman and Heggestad (1997) make a very important distinction between intelligence-as-typical performance and intelligence-as-maximal performance. Typical performance refers to the usual energy that is expended in day to day activities, whereas maximal performance refers to forced energy to produce at a maximum level. The typical intellectual engagement style is a thinking disposition which is based on the typical performance paradigm. Alternatively, the notion of intelligence is usually characterized by maximal performance (Ackerman & Heggestad, 1997). Ackerman and Heggestad (1997) argue that intelligence
tests do not correlate highly with measures of advanced academic or occupational performance because of a focus on maximal performance. As people do not usually function at maximal performance, they argue that the style of typical intellectual engagement captures what is missing in how intelligence has been conceptualized. The distinction between typical and maximal performance provide another way of conceptualizing how styles or tendencies are different from capacities or computational limitations.

The Actively Open-Minded Thinking (AOT) Disposition. The AOT scale captures tendencies such as, considering evidence that goes against one's beliefs, and a willingness to consider alternative perspectives (Baron, 1985). Tendencies toward actively open-minded thinking have been found to be significantly related to performance on a number of critical thinking tasks, including syllogistic reasoning, statistical reasoning, argument evaluation (Stanovich & West, 1997). Stanovich (1999; Stanovich & West, 1997) suggested that the AOT scale is one of many dispositional scales that tap tendencies toward epistemic self-regulation and cognitive decontextualization. The AOT disposition was included because it taps another facet of dispositional tendencies that are somewhat different from the need for cognition and reflective-impulsive types of tendencies. In addition, the AOT disposition has been found to be related to reasoning performance on other important critical thinking tasks.

Summary

An important domain of reasoning, called disjunctive thinking, was explored in this study. Shafir (1994) terms disjunctive thinking as the tendency to consider all of the possible states of the world and to evaluate the consequence of each. Based on a series of studies reviewed by Shafir (1994), he argued that people are typically reluctant to think through disjunctions.
Disjunctive thinking in the present investigation was examined in the domains of decision-making and reasoning in a set of formal and informal tasks. The formal tasks were decision-making and reasoning problems that have been studied previously in the literature, such as the Prisoner’s Dilemma (Campbell & Sowden, 1985), a decision-making problem, and the Wason Selection Task (Evans, Newstead, & Byrne, 1993), a reasoning problem. Two informal tasks were added in order to study disjunctive thinking in more realistic, everyday types of reasoning situations. The informal decision-making task was called The Consumer Task, and the informal reasoning task was called The Argument Generation Task. Based on Shafir’s (1994) definitive review, it was expected that participants would display the disjunction effect on the formal and informal tasks used in this study.

The first major goal of this study was to study disjunctive thinking problems from the decision-making and reasoning literatures, and to examine whether the problems discussed in these two separate literatures are related empirically. The first expectation was that a tendency to reason disjunctively would lead to successful performance on both the decision-making and reasoning problems. Namely, it was expected that those who reasoned disjunctively on some tasks would be more likely to reason disjunctively on the other tasks. This was called the domain-generality hypothesis.

The second purpose of this study was to examine performance on disjunctive thinking problems from an individual differences perspective. This exploration was based on Baron’s (1985) distinction between cognitive capacities and thinking dispositions, and Stanovich’s (1999; Stanovich & West, 2000) Generic Dual Process Framework for reasoning. According to Stanovich’s (1999; Stanovich & West, 2000) Generic Dual Process Framework,
cognitive ability is associated with the algorithmic level of analysis, and therefore cognitive ability tends to be less malleable than thinking dispositions. Thinking dispositions, on the other hand, are associated with the epistemic level in the Generic Dual Process Framework, and are more amenable to remediation. It was expected that the specific cognitive capacities and thinking dispositions measured in this study would explain performance on the the set of disjunctive thinking problems taken the from the decision-making and reasoning literatures. It was of particular interest to further implicate the role of thinking dispositions in disjunctive thinking performance. Thinking dispositions have received less attention in the reasoning literature, and effects involving thinking dispositions would implicate the role of processes more amenable to remediation and change in goal attainment. The key thinking dispositions examined in this study were: reflectivity and impulsivity, the need for cognition style, and the actively open-minded thinking style. The analytic strategy and tasks that were used in this study are presented in Figure 3.

Method

Participants

The participants were 128 undergraduate students (48 males and 80 females) recruited through poster advertisement on the campus of a large-sized university in Toronto, Canada.

Outliers. Univariate outliers were identified using the Mahalanobis distance metric on the cognitive ability measures. With \( p<.001 \) criterion, there were three true outliers. Two of the outliers had a very low verbal IQ score, and they also reported that English was their second language (as measured by the Vocabulary Subtest of the WAIS-R). The third outlier had a Full Scale
Figure 3. Visual Schematic of Research Strategy and Relationships of Interest

Disjunctive Thinking

Cognitive Capacities
- WAIS-R Vocabulary
- WAIS-R Block Design
- Raven's Standard Progressive Matrices
- Vocabulary Test

Demographic Variables

Thinking Dispositions
- Reflectivity-Impulsivity
- Need for Cognition Disposition
- Actively Open-Minded Thinking
- Other Exploratory Dispositions

Decision-Making Tasks
- Prisoner's Dilemma
- Newcomb's Problem
- Disease Framing Problem
- Box Problem

Reasoning Tasks
- The Selection Task
- Knights and Knaves Problem
- Charles in Scotland Problem
- Married Problem
- Green Levels Problem

Decision-Making Task
- The Consumer Task

Reasoning Task
- The Argument Generation Task
Disjunctive Thinking

IQ score of less than 80, which is below the Average range of performance (as measured by the Vocabulary and Block Design Subtests of the WAIS-R). The three outliers were removed from the analyses, as they are not part of the population with which this study is concerned (Tabachnick & Fidell, 1989). This study was concerned with obtaining participants who have at least average fluency in the English language and who have at least average intellectual ability.

There were four additional outliers in this sample; these were individuals who displayed extreme performance on individual tasks, but were included in the statistical analyses. One outlier had a very low score on Raven’s Standard Progressive Matrices, one outlier had a very low score on the Block Design Subtest of the WAIS-R, one outlier spent double the time to respond to items on the MFFT than the next participant with the second highest latency, and the final outlier had a significantly low score when the composite ability score was computed by adding the verbal and nonverbal measures of ability. These participants are sampled from the population of interest, as they performed within the average range on the verbal ability measures and on at least one of the nonverbal ability measures. Tabachnick and Fidell’s (1989) most conservative score changing option was selected for only these metrics which these participants deviated on so extremely. For each participant, the next most extreme score in the distribution was assigned to these participants so that they would deviate less, therefore minimizing the skew they created in the sample.

All but four of the participants were students at the university in which the study was conducted. Two of these participants were teachers, one who studied geography and the other who studied biology as undergraduates. The third participant had completed a medicine and a business degree, and the
fourth graduate completed an engineering degree.

The average age of the participants was 22.6 (SD=5.5) -- however, this was skewed because of a small group of mature students and university graduates. The modal age was 20 years, with a range of 18 to 49 years.

Poster advertisements were placed in many locations of the university campus, and students from a variety of diverse academic backgrounds were recruited. Fifty-two (40.6%) students were enrolled in a social science program, 32 (25%) students were enrolled in a science or math program, 30 (23.4%) students were enrolled in an arts program, 5 (3.9%) students were enrolled in a business-related program, 7 (5.5%) were enrolled in an engineering program, and 2 (1.5%) were enrolled in some other program (linguistics and speech pathology).

In the poster recruitment, participants were informed that they would be involved in a study concerning human reasoning and decision-making. They were also informed that the study would take approximately 3 hours and 15 minutes to complete, and that they would receive $20 for participation.

Decision-Making and Reasoning Tasks

Two sets of disjunctive thinking tasks were employed in this study: decision-making and reasoning tasks. These tasks are described below.

Formal Decision Making Tasks

The following formal decision-making problems were employed: Prisoner's Dilemma, Newcomb's Problem, Disease Framing Problem, and the Box Problem. An informal decision-making task, which more closely characterizes real-life decision-making, called The Consumer Task, was also employed. The Consumer Task is presented in Appendix A.

The Prisoner's Dilemma. In the classic version of this problem (as described by Shafir, 1994), two players must choose to either cooperate or
compete with the other player while being blind to the other's choice. The version of the problem used in this study was as follows:

This game was originally designed to be played by pairs of students who were sitting in front of different computers on the same computer system. Since we are not using computers today, please use your imagination, and pretend that you are sitting in front of a computer and playing the Intercollegiate Computer Game with another student.

In this game you will be presented with a situation involving you and one other player who is sitting at a computer in another room. You cannot communicate with each other. The situation requires that you make a strategic decision: to cooperate or to compete with the other player. The other player will have to make the same decision.

The situation is represented by a payoff matrix that will determine how much money each of you earns depending on whether you compete or cooperate. The matrix looks like the following:

<table>
<thead>
<tr>
<th></th>
<th>Other Cooperates</th>
<th>Other Competes</th>
</tr>
</thead>
<tbody>
<tr>
<td>You Cooperate</td>
<td>You: 20</td>
<td>You: 5</td>
</tr>
<tr>
<td></td>
<td>Other: 20</td>
<td>Other: 25</td>
</tr>
<tr>
<td>You Compete</td>
<td>You: 25</td>
<td>You: 10</td>
</tr>
<tr>
<td></td>
<td>Other: 5</td>
<td>Other: 10</td>
</tr>
</tbody>
</table>

According to this matrix, if you both cooperate you will both earn $20 each. If you cooperate and the other person competes, the other will earn $25 and you will earn only $5. Similarly, if you compete and the other person cooperates, you will earn $25 and the other person will earn only $5. Finally, if you both choose to compete, you will each earn $10. Not knowing what the other person will choose to do, what would you choose?

(a) I would choose to compete
(b) I would choose to cooperate

If both players decide to cooperate, they each gain $20. If both decide to
compete, they each gain $10. However, if one competes and the other cooperates, then the competitor gains $25 and the cooperator gains $5. Shafir (1994) argues that because neither player knows what the other will select, the best option for each player is to compete rather than to cooperate in order to maximize individual gain. That is, if one considers the possible outcome for each alternative, it becomes evident that $25 (you compete other cooperates) is greater than $20 (you both cooperate) and $10 (you both compete) is greater than $5 (other competes you cooperate). Shafir (1994) argues that people's choices are affected by conflicting motivations, and the preferred outcome results from mutual cooperation instead of competition. However, disjunctive consideration of each possible outcome helps the player realize that the optimal strategy is to compete in order to achieve maximal gain. This single play game described here was employed in the present study in which participants had to decide whether they would Compete and Cooperate. It was found that 54% (n=66) of participants chose the “compete” response and 46% (n=57) of participants chose the “cooperate” response.

**Newcomb's Problem.** Originally discussed by Nozick (1969), Shafir (1994) describes Newcomb's Problem as an instance of “quasi-magical thinking”. The problem is as follows:

![Box A: $20, Box B: $250 or $0?](image)

Here is another problem that asks you to make use of your imagination. Consider the two boxes above. Box A contains $20 for sure. Box B may or may not contain $250. Pretend your options will be to: [Don’t choose yet]

1. Choose both boxes (and collect the money that is in both).
2. Choose Box B only (and collect only the money that is in Box B).
Imagine that we have a computer program called the "Predictor" that has analyzed the pattern of the responses you have already made to all of the earlier questions. Based on this analysis, the program has already predicted your preference for this problem and has already loaded the boxes accordingly. If, based on this analysis of your previous preferences, the program has predicted that you will take both boxes, then it has left Box B empty. On the other hand, if it has predicted that you will take only Box B, then it has already put $250 in that box. So far, the program has been very successful: Most of the participants who choose Box B received $250; in contrast, few of those who chose both boxes found $250 in Box B.

Which of the above options would you choose?

(1) Choose both boxes
(2) Choose Box B only

Shafir (1994) suggests that disjunctive consideration in this problem is required for maximal gain. For maximal gain, the best option is to select Both Boxes rather than selecting Box B Only, as there is the potential of gaining a total of 270 points (20 points in Box A and 250 points in Box B). This version of Newcomb's Problem was used in this study, and only 37% (n=46) of participants chose the Both Boxes response, and 63.2% (n=79) of participants chose the Box B Only response.

The Disease Framing Problem. Originally from Tversky and Kahneman (1981), the Disease Framing Problem is composed of two parts. The two parts of the problem are as follows:

First Presentation: Imagine that the United States is preparing for the outbreak of an unusual Asian disease expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

a) If Program A is adopted, 200 people will be saved.

b) If Program B is adopted, there is a \( \frac{1}{3} \) probability that 600 people will be saved, and a \( \frac{2}{3} \) probability that no people will be saved.

If it was your decision to choose between Program A or Program B, which program would you choose? (circle a) or b)
Second Presentation:
Imagine that the United States is preparing for the outbreak of an unusual Asian disease expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

a) If Program C is adopted, 400 people will die.
b) If Program D is adopted, there is a 1/3 probability that nobody will die, and a 2/3 probability that 600 people will die.

If it was your decision to choose between Program C or Program D, which program would you choose? (circle a) or b))

The two parts of this problem (First Presentation and Second Presentation) were presented on separate occasions in the testing battery. The difference between the two programs in each problem is that one program presents the rate of lives saved/lives lost in absolute numbers, and the other program presents the rate of lives saved/lives lost in terms of probabilities. The framing effect occurs when participants are inconsistent in their program preference in each version of the problem -- when in reality, the programs are directly equivalent. The reader should refer to the introduction for a more detailed discussion of this problem.

In the present study, the lives saved version was presented first to all participants, and the lives lost version was presented secondly. Three different response tendencies were coded for this problem:

- framing effect: the participant makes an inconsistent selection in order to select options worded with a more favourable outcome; ‘200 saved’ and ‘2/3 lives lost’;

- no framing effect: participant selects consistently; ‘200 saved’ and ‘400 lives lost’ OR ‘1/3 saved’ and ‘2/3 lives lost’; and

- reverse frame: participant makes an inconsistent selection in order to select options worded with a less favourable outcome; ‘2/3 lives lost’ and ‘400 lives lost’.

A minority of participants, 3.2% (4 participants), displayed a reverse
framing effect. There is no compelling theoretical rationale for a reverse framing effect, because participants are choosing the riskier option for lives saved and the absolute value for lives lost. Using a charitable analysis, this response pattern may reflect an indifference between the options. However, it is also possible that these participants were providing a random response pattern. As only four participants demonstrated a reverse framing effect, these participants were excluded from the analyses. Therefore, those who displayed a framing effect and those who displayed no framing effect were included in the statistical analyses.

The Box Problem. This problem of probabilistic judgment was first studied by Shafir (1994), and in this problem, participants are presented with the following choice:

Imagine that in front of you are two boxes. Inside each of the boxes is a ball that is equally likely to be either white, blue, or purple.

![Box A and Box B](image)

You are now offered to play one of the following two games of chance:

**Game 1:** You guess the color of the ball in the left-hand box. You win 50 dollars if you were right, and nothing if you were wrong.

**Game 2:** You choose to open both boxes. You win 50 dollars if the balls in the boxes are the same color, and nothing if they are different colors.

Which would you prefer to play?

a) Game 1  
b) Game 2  
c) No Preference

In this problem, if one disjunctively considers the outcome of each game, it
becomes evident that there is a 33% chance of winning either game. Therefore, participants should be indifferent about which game they choose to play by selecting the “no preference” option. On this version of the Box Problem, 24% (n=30) of participants selected the “no preference” response. Then, 67% (n=84) of participants selected the Game 1 option, and 9% (n=11) selected the Game 2 option.

Informal Decision Making Task

The Consumer Task. This task was based on the methodology of Verplanken (1993). In the task used here, called The Consumer Task, participants were asked to list those features that they would consider before purchasing a bread maker, a t-shirt, and a car. These particular items were selected because they provided a reasonable range in: cost to purchase, functionality/practicality, and number of available features to consider.

Two scoring approaches were used. First, a scoring key was developed based on some pilot research in order to code for major categories. The pilot study was conducted with 19 participants who attended a local city college, and they participated for course credit. The purpose of this scoring key was to a priori have an exhaustive set of possible features and to eliminate redundant features. This scoring guide is presented in Appendix B. Participants' score based on this scoring guide reflected the total number of non-redundant features generated for each product. Second, the May 1997 issue of Consumer Reports magazine was used to obtain a list of features that consumer reports studies have identified as diagnostic features that differentiate different brands of products. Each participant received a score out of three for the bread maker, a score out of four for the t-shirt, and a score out of four for the car.

Participants were also asked to report whether they have previously purchased the product, and if so, how many times. This brief questionnaire
Disjunctive Thinking

was presented prior to presentation of the Consumer Task. The questions were as follows:

1) Have you ever purchased a bread maker or helped someone select a bread maker to purchase before?
   Yes____ No____ If Yes, estimate how many times in total: ______

2) Have you ever purchased a t-shirt or helped someone choose a t-shirt to purchase before?
   Yes ____ No____ If Yes, estimate how many times in total: ______

3) Have you ever purchased a car or helped someone choose a car to purchase before?
   Yes ____ No____ If Yes, estimate how many times in total: ______

These questions were added to consider the role of experience, because presumably someone who has purchased a car before may generate more features, for example, than someone who has never purchased a car.

**Formal Reasoning Problems**

A series of formal reasoning problems that have typically been studied in the area of cognitive psychology were employed. These problems include: the Selection Task (Wason, 1966), the Knights and Knaves Problem (Rips, 1989; Shafir, 1994), the Charles in Scotland Problem (Baron, 1995), and the Married Problem and the Green Levels Problem (both from Levesque, 1986; 1989). One additional informal reasoning task was also employed, with the same goals in mind as for the Consumer Task: to more closely mimic real-life situations in which disjunctive consideration of alternatives leads to better thinking. This task, called The Argument Generation Task, is presented in Appendix C.

**The Selection Task.** Originally used by Wason (1966), the abstract version of the Selection Task is a reasoning problem that has been studied extensively in the deductive reasoning literature (e.g., Evans & Over, 1996;
Disjunctive Thinking

Newstead & Evans, 1995; Evans, Newstead, & Byrne, 1993). This problem has been studied as an instance of hypothesis testing and deductive reasoning based on the logic of conditionals, but it also involves disjunctive consideration as one has to consider the consequences of turning each card for successful performance (Evans, Newstead, & Byrne, 1993).

In the abstract version of this problem, participants are presented with four cards, as follows:

Each of the boxes below represents a card lying on a table. Each one of the cards has a letter on one side and a number on the other side. Here is the rule: If a card has a vowel on its letter side, then it has an even number on its number side. As you can see, two of the cards are letter-side up, and two of the cards are number-side up. Your task is to decide which card or cards must be turned over in order to find out whether or not the rule is being violated. Indicate which card or cards must be turned over by placing check marks in the appropriate places.

<table>
<thead>
<tr>
<th>K</th>
<th>A</th>
<th>8</th>
<th>5</th>
</tr>
</thead>
</table>

Turn over?  Turn over?  Turn over?  Turn over?
yes_______  yes_______  yes_______  yes_______
no_______  no_______  no_______  no_______

The participants' task is to determine which of the four cards should be turned over in order to determine if the rule has been violated. It is typical for participants to immediately select the cards mentioned in the rule, that is, to select the vowel and the even number -- this has been termed the matching bias (Evans, Newstead, & Byrne, 1993). However, if participants consider what letter or number could be on the other side of each card, then they are likely to realize which cards would necessarily break the rule. This exhaustive consideration is the disjunctive component of this task. The cards that should be turned over are the card with a vowel on one side (termed "P") and the card
with an odd number on one side (termed "Not Q").

Overall, 7.2% gave the correct P, Not-Q response. No one selected P,Q, Not-Q, 16.8% selected P only, and 22.4% selected all four cards ('All'). The classic matching response, P,Q, was selected by 40% of participants. Then, 13.6% of participants selected other incorrect responses. The other incorrect responses included: five participants who chose Not-P, Not-Q; four participants who chose Not-P, Q; two participants who chose Q only; two participants who chose Not-P, P; two participants chose none of the cards, one participant chose Q, Not-Q; and one participant who chose Not-P, Q, Not-Q.

For statistical parsimony and based on a theoretical rationale (see Results section), some composites were formed for the statistical analyses. Those who provided the correct P, Not-Q response, the P only response, and All of the cards were collapsed (Cor+P+All) to form the higher level, analytic responders group. The comparison group were the P,Q responders, who were the non-analytic, heuristic comparison. Those who provided other incorrect responses were eliminated from further analyses because it was unclear whether their responses were disjunctive or non-disjunctive.

Knights and Knaves Problem. Originally from Smullyan (1978), this problem has been studied in the classic deductive reasoning literature (such as Rips, 1989; Johnson-Laird & Byrne, 1990). The problem is as follows:

Imagine that there are three inhabitants of a fictitious country, A, B, and C, each of whom is a either a knight or a knave. Knights always tell the truth. Knaves always lie. Two people are said to be of the same type if they are both knights or both knaves. A and B make the following statements:

A says: "B is a knave"
B says: "A and C are of the same type"

What is C? a) a Knight b) a Knave c) Cannot be Determined

The reader should refer to the introduction for a detailed discussion of this
problem. In this problem, 9.6% (n=12) of participants selected the “knight” response option, 41.6% of participants selected the “knave” response, and 48.8% of participants selected the “cannot be determined” response. There is no theoretical rationale for a knight response selection, and as there were only a minority of these responses, knight responders were eliminated from the analyses.

**Charles in Scotland Problem.** This problem is also taken from the deductive logic literature (Johnson-Laird, Byrne, & Schaeken, 1992), and has been discussed in some detail by Baron (1994). The problem is as follows:

Consider the following two facts:

June is in Wales or Charles is in Scotland, but not both.
Charles is in Scotland or Kate is in Ireland, but not both.

What, if anything, follows from these two facts?

If one works through the possible combinations of who can where at what time, the conclusion we can draw is that:

\[
\text{EITHER (Charles is in Scotland) OR (BOTH June is in Wales AND Kate is in Ireland), but not both. (1)}
\]

The correct solution can only be derived by elaborating the full model, that is, by considering both premises and the possible locations of each person. Therefore, a response which reflected elaboration of the full model, as in (1), was taken as an index of disjunctive consideration of the alternatives.

After the data was collected, seven different response categories were coded for this problem, including:

1. Correct (‘Charles is in Scotland OR June is in Wales and Kate is in Ireland.’);
2. June/Kate (‘June is in Wales and Kate is in Ireland.’);
3. Partial Contingency (‘If Charles is in Scotland, then June is not in
Wales, nor is Kate in Ireland.

4. Nothing ('Nothing');
5. Blank (space left blank);
6. Charles is in Scotland ('Charles is in Scotland');
7. Other Incorrect Responses

A complete list of all of the variations given by participants and how each variation was coded is presented in Appendix D.

In this sample, only 12.8% participants provided the fully disjunctive correct response. Then, 21.6% of participants responded with 'Charles is in Scotland', 24% provided an incorrect response ('Other Incorrect'), 8.8% responded 'nothing', and 8.8% left the question blank. There were two other patterns of responses which were of interest: 'partial contingency' responders and 'June/Kate' responders. For statistical parsimony and based on a theoretical rationale (see Results section), some composites were formed for the statistical analyses. Consequently, three different sets of comparisons were performed:

1. 'Correct' and 'Partial Contingency' (Cor+Partial), also called the highly analytic group, were compared to all of the other responses (June/Kate+Charles+Nothing+Blank+Other);

2. 'Correct', 'Partial Contingency', and 'June/Kate' (Cor+Partial+June/Kate), also called the broadly analytic group, were compared to the other responses (Charles+Nothing+Blank+Other); and

3. 'Correct', 'Partial Contingency', and 'June/Kate' (Cor+Partial+June/Kate) were compared to 'Charles is in Scotland' responders (Charles). This was the sharpest contrast -- between the broadly analytic group and the least analytic or heuristic responders.

The Married Problem. From Levesque (1986), the problem is as follows:

Jack is looking at Ann but Ann is looking at George. Jack is married but George is not. Is a married person looking at an unmarried person?

A) Yes  B) No  C) Cannot be determined

What Levesque defines as implicit information in this problem is analogous to
what has been called disjunctive consideration in the present study. That is, one needs to consider the disjuncts of Ann's marital status, which is the implicit information, in order to derive the correct conclusion. If Ann is married, then the answer is yes because she would be looking at George who is unmarried. If Ann is not married, then the answer is still yes because Jack, who is married, would be looking at Ann. The correct solution to this problem can only be derived by using a disjunctive strategy. On this problem, 1.6% (n=2) of participants selected the 'no' response alternative, 12.8% (n=16) of participants selected the 'yes' response, and 85.6% (n=107) of participants selected the 'cannot be determined' response. It is difficult to infer how 'no' responders arrived at their solution, and as only two participants selected this option, they were eliminated from the analyses.

**The Green Levels Problem.** Also from Levesque (1989), the problem is:

There are 5 blocks in a stack, where the second one from the top is green, and the fourth is not green. Is there a green block directly on top of a non-green block?

A) Yes  B) No  C) Cannot be determined

Similar to the Married Problem, Levesque (1989) discusses the Green Levels Problem as an instance in which important information to solve the problem is not explicit in the premises, but needs to be inferred from the premises. One needs to consider the colour of the third block in the stack, which is the
implicit information in this problem. If the third block is green, then the answer is yes because it sits directly over the fourth block which is not green. If the third block is not green, then the answer is still yes because it sits directly under the second block which is green. On this problem, 7.2% (n=9) of participants selected the 'no' response, 8.8% (n=11) of participants selected the 'yes' response, and 84% (n=105) of participants selected the 'cannot be determined' response alternative. To be consistent with the analyses of the Married Problem, the low frequency 'no' responders were eliminated from the analyses.

Informal Reasoning Task

The Argument Generation Task. This task was based on the methodology used by Perkins, Farady, and Bushey (1991). In the original task used by Perkins et al. (1991), participants were presented with a question about a current issue, and they were asked to generate arguments both for and against their position on the issue. The current issue used by Perkins et al. (1991) was whether more funding would improve education. Perkins et al. (1991) scored the number of arguments generated in favour and the number of arguments generated against each participants' position on the issue.

In the current study, participants were presented with three different issues. These issues were selected in collaboration with a research group composed of six members. The group generated a set of issues that were controversial, thought-provoking, and that had a number of potential arguments available both for and against the issue. The following three issues were selected and used in this study:

1. The real cost of a university education is $12000/year. Students are currently paying approximately $3500 in tuition. The difference is paid by the taxpayer. University students should pay for the full cost of their university education.

2. People should be allowed to sell their organs.
3. The cost of gasoline should be doubled to discourage people from driving.

Participants were first asked to indicate their position on each of these three issues. This prior belief measure was obtained in order to identify which arguments endorsed the participants’ position (my-side arguments) and which arguments did not endorse the participants’ position (other-side arguments) for each issue. Participants were asked to indicate their position on each issue using the following scale:

Indicate to what extent you agree or disagree with this statement. Please circle one of the following:

1 - Disagree Strongly
2 - Disagree Moderately
3 - Disagree Slightly
4 - Agree Slightly
5 - Agree Moderately
6 - Agree Strongly

After being administered several other unrelated reasoning tasks, participants were once again presented with the same three issues. This time their task was to generate arguments both for and against their own position on each issue. Of interest was the total number of arguments that participants generated which endorsed their beliefs (my-side arguments), and the total number of arguments that do not endorse their beliefs (other-side arguments). Some pilot work was done in order to obtain a suitable and scorable structure for this task.

A coding scheme was developed from previous pilot data in which the tuition and gasoline issues were utilized in a similar paradigm. The two main categories of interest were supportive and refutational arguments, because these two categories provided the basis for coding the number of my-side and other-side arguments. Supportive arguments were defined as those
arguments which provided information that supported the issue. For example, an argument endorsing the tuition issue was: "Should taxpayers have to foot bill for those who they don't even know?" Refutational arguments included those arguments which provided information that refuted the issue. An example of a refutational argument for the tuition issue was: "Many students would simply not be able to attend college."

A number of other categories were also coded in order to more succinctly capture the different types of arguments that were given. These categories included: ambiguous arguments, irrelevant arguments, statements of lack of knowledge, reiterations of supportive or refutational arguments, and opinions with a supportive or refutational valence. Ambiguous arguments were defined as those arguments which were ambiguous in terms of the position taken on the issue or they were simply uncodeable for some reason. An example for the tuition issue was: "Education is overpriced." Irrelevant arguments were those arguments which were not apparently related to the issue that was given to participants. For example, an irrelevant argument for the tuition issue was: "If parents can afford it, they should help." In this case, the participant did not address the cost of tuition itself, rather described who should help to fund the student. The lack of knowledge arguments included those in which participants stated that they knew very little about the issue. An example for the tuition issue was: "I don't know enough about the facts and figures to really make a legitimate argument." A reiteration was coded for redundant supportive or redundant refutational arguments. That is, if the same participant gave a supportive (or refutational) argument that was a re-statement of a previous argument for the same issue, then the argument was coded in this category. For example, a participant gave the two following supportive arguments for the gasoline
issue: "The idea is nice, but it might be horrible for people who can't afford it." and "I agree with statement, lots of people could be screwed." Opinions were coded as those statements which only provided an opinion and no argumentative content, and opinions were coded as having a supportive or refutational valence. A supportive opinion, for example, was the following for the gasoline issue: "Good point." An opinion with a refutational valence for the tuition issue was: "I don't agree with this."

Ninety-eight percent of the arguments were coded as either supportive or refutational, while 2% were coded in the other categories. There were a total of 1986 arguments (three issues for each of 128 participants). Two coders used a scoring key composed of samples of each argument type. The coders coded all of the arguments independently and they were blind to the participant's prior belief on the issues. There was 91% agreement between the two coders. Discrepancies were resolved by having both coders discuss scoring differences and determine the most appropriate coding. Arguments in which the coders could not easily resolve the discrepancies were passed onto a third coder. There were a total of 30 arguments that required resolution by the third coder.

The Argument Generation Task and the Role of Knowledge.
Participants were asked some questions to test their general knowledge on the issues of tuition and gasoline. The purpose of these questions was to consider the role of knowledge in the number of and types of arguments generated. More specifically, if people's beliefs are driven by their knowledge about the issue, and if these same people generate more my-side than other-side arguments, then this is qualitatively different than those who lack knowledge and generate more my-side arguments. These investigations were exploratory in nature.

Participant's domain knowledge was tested by asking specific questions
which were compiled from Statistics Canada and other reputable government publications. Some pilot work was done with students at a local community college in order to eliminate questions that were either too difficult or those that everyone responded to correctly. The final set of questions included:

**Tuition Issue**

1. Currently students are paying $3769 (with incidental fees) in tuition at the University of Toronto. The University of Michigan is a comparable institution. How much tuition do you think in-state students at the University of Michigan are paying in tuition? 
Correct Solution: approximately $10000-$12000/year Canadian

2. How much do you think young people (between 20-29 years of age) earn on average per year, if:
   a) they have a Doctorate degree?
   Correct Solution: approximately $45000/year Canadian

**Gasoline Issue**

1. Currently, Canadians are paying between $0.56-$0.60 cents/litre for gasoline. How much do you think people are currently paying for gasoline in other countries per litre in Canadian funds?
   a) Japan $____/litre  Correct Solution: approximately $1.49 Canadian
   b) Italy $____/litre  Correct Solution: approximately $1.62 Canadian
   c) Germany $____/litre  Correct Solution: approximately $1.59 Canadian
   d) England $____/litre  Correct Solution: approximately $1.20 Canadian

In order to capture knowledge in the most general sense, the knowledge questions were coded by creating the broad categories of ‘correct’ response and ‘misconception.’ The tuition in Michigan question and the salaries of people with doctorates question were used for the tuition issue. The cost of gasoline in other countries question was used for the gasoline issue. The purpose of the scoring procedures was to capture very general knowledge, namely that tuition at Michigan State University was more expensive than in Canadian universities, that a doctorate degree does not translate into a very
high salary upon graduation, and that gasoline in other parts of the world, such as Europe, is considerably more expensive than the cost of gasoline in Canada. It was expected that knowledge would further discriminate between better and poorly calibrated reasoners.

For the tuition in Michigan question, 30.65% (n=38; range $2000-$5000) of participants were coded as displaying a ‘misconception’, and 69.35% (n=86; range $5200-$30000) of participants were coded as displaying the ‘correct’ response. For the salaries of recent doctorates, 24.59% (n=30; range $80000-$800000) of participants were coded as displaying a ‘misconception’, and 75.41% (n=92; range $20000-$70000) of participants were coded as displaying the ‘correct’ response. For the cost of gasoline in other countries question, 55.65% (n=69; range $0.475-$0.975) of participants were coded as displaying a ‘misconception’, and 44.35% (n=55; range $1.00-$35.00) of participants were coded as displaying the ‘correct’ response.

**Thinking Dispositions**

The Thinking Dispositions were measured using two main tasks. One task was the Matching Familiar Figures Test (MFFT; a sample item is displayed in Appendix E), and the second task was the Thinking Dispositions Questionnaire (a complete set of the items appears in Appendix F).

**The Matching Familiar Figures Test (MFFT)**

The Matching Familiar Figures Test (MFFT) was also used to measure the dimension of reflectivity and impulsivity. The MFFT version developed by Kagan, Rosman, Day, Albert, and Phillips (1964) was employed in the present study. In this task, participants were presented with a target picture of an object, and their task was to find the correct match from an array of six other pictures. Participants’ latency and number of errors were measured for each choice and for each item. When participants made an incorrect selection, they
were asked to select again -- this was repeated until the participant found the correct match (up to a maximum of six possible responses). Participants who were above the median on MFFT response time and below the median on errors were defined as "reflective", while participants who were below the median on MFFT response time and above the median on errors were defined as "impulsive" (Messer, 1976).

The mean time to the first response for all items and the total number of errors for all items were standardized for each participant. The standardized error metric was called MFFT\textsubscript{Errors} and the standardized metric for reaction time was called MFFT\textsubscript{RT}. Then the difference between these standard scores was taken to create a variable which took into account both response time and number of errors. This variable was called MFFT\textsubscript{RT-Errors}.

**Thinking Dispositions Questionnaire**

Participants completed a questionnaire consisting of a number of self-report subscales. The response format for each item in the questionnaire was: Strongly Agree (6), Moderately Agree (5), Slightly Agree (4), Slightly Disagree (3), Moderately Disagree (2), and Strongly Disagree (1). The items from the different subscales were randomly intermixed in the questionnaire. A brief description with some examples of these subscales follow. Two main thinking disposition composites were formed based on these subscales, and other related dispositions were used for further exploratory analyses. The two composite thinking dispositions were the Need for Cognition Composite and the Actively Open-Minded Thinking Composite (from Stanovich and West, 1997).

**The Need for Cognition Composite.** The Need for Cognition Composite was formed by summing the standard scores of the total scores on the Need for Cognition Scale and the Typical Intellectual Engagement scales.
The most recent Need for Cognition Scale was utilized in this study. This self-report scale was composed of 18 items taken from Cacioppo, Petty, Feinstein, and Jarvis (1996). Sample items include: "The notion of thinking abstractly is appealing to me", and "I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought". A subset of 18 items from the Typical Intellectual Engagement (TIE) Scale was selected for use in this study from the 56-item scale developed and employed by Goff and Ackerman (1992), and Ackerman and Heggestad (1997). Only a subset of the original scale was used, as many of the items were redundant within the scale and with items on the Need for Cognition Scale. Some of the items from the TIE include: "You enjoy thinking out complicated problems", and "I would enjoy hearing details about discoveries in any field".

The Actively Open-Minded Thinking (AOT) Composite. The AOT Composite is formed from the standard scores of the total scores on the following scales: the AOT scale, the NEO-Ideas subscale, the Dogmatism subscale, and the Categorical Thinking Scale. The first two scales are summed and the latter two are subtracted from the composite scale.

There were ten items on the AOT scale, some tapping the disposition toward reflectivity (e.g., "If I think longer about a problem I will be more likely to solve it"), willingness to consider evidence contradictory to beliefs (e.g., "People should always take into consideration evidence that goes against their beliefs"), willingness to consider alternative opinions and explanations ("A person should always consider new possibilities"), and a tolerance for ambiguity combined with a willingness to postpone closure ("Changing your mind is a sign of weakness" -- which is reverse scored).

From the Revised NEO Personality Inventory (Piedment, 1998; Costa &
McCrae, 1992), the eight items from the Openness-Ideas facet were
administered (e.g., “I have a lot of intellectual curiosity”, “I find philosophical
arguments boring”--the latter reverse scored).

The Dogmatism Subscale consisted of three items taken from a short-
form field version (Troldahl, & Powell, 1965) of Rokeach’s (1960) dogmatism
scale. The three items were “Of all the different philosophies which exist in
the world there is probably only one which is correct”; “Even though freedom of
speech for all groups is a worthwhile goal, it is unfortunately necessary to
restrict the freedom of certain political groups”; and “There are two kinds of
people in this world: those who are for the truth and those who are against the
truth”.

Three items from the Categorical Thinking Subscale of Epstein and
Meier’s (1989) Constructive Thinking Inventory (CTI) were administered:
“There are basically two kinds of people in this world, good and bad”; “I think
there are many wrong ways, but only one right way, to almost anything”; “I
tend to classify people as either for me or against me.”

Other Dispositions. The 11 item Objectivism Scale was used here and
was developed by Leary, Shepperd, McNeil, Jenkins, and Barnes (1986). This
scales includes items, such as “I seek as much information as possible before
making decisions”, and ”I think the answers to most questions in life can be
found through careful, objective analysis of the situation”.

The seven item Vigilance subscale was taken from the Melbourne
Decision-Making Questionnaire developed by Mann et al. (1997). The
questionnaire developed by Mann et al. (1997) was designed to measure
tendencies and coping patterns in conflict theory of decision-making. The
Vigilance subscale is made up of items, such as ”I like to consider all of the
alternatives”, and ”I try to find out the disadvantages of all alternatives".
The eight items from the Openness-Values facet of the Revised NEO Personality Inventory (Piedment, 1998; Costa & McCrae, 1992) were also administered. Sample items included: “I believe that laws and social policies should change to reflect the needs of a changing world”, and “I believe letting students hear controversial speakers can only confuse and mislead them”—the latter reverse scored).

Participants answered one item taken from Paulhus and Reid (1991) as a measure of belief in prayer: “In one way or another, God answers all my prayers”. Participants also indicated their level of religiosity on the following item from Hunsberger, Lea, Pancer, Pratt and McKenzie (1992): “Overall, I would say that I am a religious person”.

A paranormal beliefs subscale was used and it was composed of six items. Two items were concerned with belief in astrology (“It is advisable to consult your horoscope daily”) and were adapted from the paranormal belief scale validated by Jones, Russell, and Nickel (1977). The four remaining items concerned the belief in the concept of luck (e.g., “I have personal possessions that bring me luck at times”) and were similar to items on the superstition subscale of a paranormal beliefs questionnaire developed by Tobacyk and Milford (1983).

Five items reflecting social desirability response bias were taken from (Paulhus, 1991). Some examples of these items are “I always obey laws, even if I’m unlikely to get caught”, and “I sometimes tell lies if I have to”— the latter reverse scored.

**General Ability Measures**

Participants completed a short form of the Wechsler Adult Intelligence Scale-Revised (WAIS-R, Wechsler, 1974). The short form consisted of the Vocabulary (verbal measure) and Block Design (nonverbal measure) subtests.
Sattler (1992) reports that the combination of the Block Design and the Vocabulary subtests provide the most reliable two-subtest estimate of Full-scale IQ, where reliability equals .90. Instead of prorating the two subscales in order to obtain an estimated Full-scale IQ score, the total raw scores on the Vocabulary and Block Design were standardized. This method was chosen because our purpose was to derive a fixed measure of cognitive ability without any corrections for age.

In addition to the WAIS-R subtests, another verbal and nonverbal measure were also used. The verbal measure was a brief vocabulary measure employing the checklist-with-foils format that has been shown to be a reliable and valid way of assessing individual differences in vocabulary knowledge (Anderson & Freebody, 1983; Cooksey & Freebody, 1987; Zimmerman, Broder, Shaughnessy, & Underwood, 1977). The stimuli for the task were 40 words and 20 pronounceable nonwords taken largely from the stimulus list of Zimmerman et al. (1977). The words and nonwords were intermixed via alphabetization. A copy of the Vocabulary Checklist appears in Appendix G. The subjects were told that some of the letter strings were actual words and that others were not and that their task was to read through the list of items and to put a check mark next to those that they knew were words. Scoring on the task was determined by taking the proportion of the target items that were checked and subtracting the proportion of foils checked. Other corrections for guessing and differential criterion effects have also been previously used (see Snodgrass & Corwin, 1988), resulting in virtually identical correlational results.

The nonverbal measure consisted 18 items from Raven’s Advanced Progressive Matrices (Set II, Raven, 1962), a task tapping analytic intelligence and that is commonly viewed as a good measure of g (Carpenter et al., 1990).
The students were given 15 minutes to complete the 18 items. The twelve easiest items were eliminated because performance in a college sample has been found to be near ceiling and six of the most difficult problems were eliminated because performance is nearly floored (Carpenter et al., 1990; Raven et al., 1977). The remaining 18 items were utilized in order to achieve a cut-time version so that the Advanced Matrices would still have adequate reliability and discriminating power. A previous investigation used a 16-item version of the Standard Progressive Matrices for cut-time administration and achieved reliabilities over .75 in samples of children (Cahan & Cohen, 1989).

The General Cognitive Ability measure was formed by summing the standardized scores of the WAIS-R Vocabulary Subtest, the WAIS-R Block Design Subtest, the Vocabulary Checklist, and the Raven's Progressive Matrices scores. The Verbal Ability measure was the composite of the WAIS-R Vocabulary Subtest and the Vocabulary Checklist. The Nonverbal Ability measure was the composite of the WAIS-R Block Design Subtest and the Raven's Progressive Matrices scores.

Procedure

Participants were all tested individually with the same experimenter. All participants read and signed a consent form prior to commencing the testing. All participants received the common test battery in the same sequential order. Appendix H lists the order of Task Administration.

At approximately half way through the testing battery, participants were asked by the experimenter if they wanted to take a break. One hundred and sixteen participants (90%) took a break. All participants were offered a soft drink or juice and a snack at the scheduled break regardless of whether they chose a break or not. The average time required to complete the test battery was 162 minutes (SD=31.55 minutes).
Results and Discussion

Four main sections are outlined in the results. First, performance is described on each of the formal decision making and reasoning problems, and on the informal decision making and reasoning tasks. In particular, frequencies of participants who displayed the disjunction effect are described, and replications of past research are considered on the formal tasks. Second, the domain generality question was addressed, both within and across the different tasks. Third, the relationships between the formal and informal tasks with cognitive ability and dispositional measures are outlined. A final section includes other relevant patterns and findings.

1. Description of Performance

An overview of performance on all of the formal problems is presented in Table 5. The two main categories of interest were the disjunctive and non-disjunctive responses. The “other responses of interest” presented in Table 4 were included in some of the analyses, and the “unusual responses” were eliminated from the analyses (see Methods for further details). Performance on each problem and task is described.

Formal Decision Making Problems

Prisoner's Dilemma. In the single trial version of the Prisoner's Dilemma used in this study, 53.7% (n=66) of participants chose to 'compete' and 46.3% (n=57) of participants chose to 'cooperate.' Shafir and Tversky (1992; also reported by Shafir, 1994) found that 37% (n=80) of their participants chose to cooperate and Stanovich (1999) reported that 41.2% (n=661) of participants chose to cooperate in analogous versions of this problem. Overall, a slightly higher percentage of participants chose to cooperate in this study, but findings are still in line with previous research.

There has been some debate about the optimal choice in the one trial
Table 4

Formal Decision-Making and Reasoning Problem Performance

<table>
<thead>
<tr>
<th>Variable</th>
<th>Disjunctive Responses</th>
<th>Non-Disjunctive Responses</th>
<th>Other Responses Of Interest</th>
<th>Unusual Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Formal Decision Making Problems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prisoner’s Dilemma</td>
<td>“Compete”</td>
<td>“Cooperate”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>53.7% (n=66)</td>
<td>46.3% (n=57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Newcomb’s Problem</td>
<td>“Both Boxes”</td>
<td>“B Only”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>36.8% (n=46)</td>
<td>63.2% (n=79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease Framing Problem</td>
<td>“No Framing”</td>
<td>“Framing”</td>
<td></td>
<td>“Reverse Framing”</td>
</tr>
<tr>
<td></td>
<td>75.6% (n=93)</td>
<td>21.1% (n=26)</td>
<td></td>
<td>3.3% (n=4)</td>
</tr>
<tr>
<td>Box Preference Problem</td>
<td>“No Preference”</td>
<td>“Game 1”</td>
<td>“Game 2”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>24.0% (n=30)</td>
<td>67.2% (n=84)</td>
<td>8.8% (n=11)</td>
<td></td>
</tr>
<tr>
<td><strong>Formal Reasoning Problems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection Task</td>
<td>“Cor+P+All”</td>
<td>“P,Q”</td>
<td></td>
<td>“Other”</td>
</tr>
<tr>
<td></td>
<td>46.4% (n=58)</td>
<td>40.0% (n=50)</td>
<td></td>
<td>13.6% (n=17)</td>
</tr>
<tr>
<td>Knights &amp; Knaves</td>
<td>“Knave”</td>
<td>“CBD”</td>
<td></td>
<td>“Knight”</td>
</tr>
<tr>
<td></td>
<td>41.6% (n=52)</td>
<td>48.8% (n=61)</td>
<td></td>
<td>9.6% (n=12)</td>
</tr>
<tr>
<td>Charles in Scotland</td>
<td>“Cor+Par+J/K”</td>
<td>“Charles in Scotland”</td>
<td></td>
<td>“Other”</td>
</tr>
<tr>
<td></td>
<td>36.8% (n=46)</td>
<td>21.6% (n=27)</td>
<td></td>
<td>41.6% (n=52)</td>
</tr>
<tr>
<td>Married Problem</td>
<td>“Yes”</td>
<td>“CBD”</td>
<td></td>
<td>“No”</td>
</tr>
<tr>
<td></td>
<td>12.8% (n=16)</td>
<td>85.6% (n=107)</td>
<td></td>
<td>1.6% (n=2)</td>
</tr>
<tr>
<td>Green Problem</td>
<td>“Yes”</td>
<td>“CBD”</td>
<td></td>
<td>“No”</td>
</tr>
<tr>
<td></td>
<td>8.8% (n=11)</td>
<td>84% (n=105)</td>
<td></td>
<td>7.2% (n=9)</td>
</tr>
</tbody>
</table>

**Note:** CBD = Cannot be Determined;  
Cor+P+All = Correct Response + P Only Choosers + All 4 Card Choosers;  
Cor+Par+J/K = Correct Response + Partial Contingency Responders + June/Kate Choosers.
version of the Prisoner's Dilemma. Some have argued that the logic in this problem parallels the logic of Newcomb's Problem (Lewis, 1979), and that the compete selection is the consequentialist, disjunctive solution. The compete response is treated as the optimal response in the analyses presented here, but the normative status of this response deserves further consideration.

**Newcomb's Problem.** In this decision making problem, only 36.8% (n=46) of participants chose 'both boxes,' and 63.2% (n=79) of participants chose 'Box B only.' In a sample of 40 participants, Shafir (1994; Shafir & Tversky, 1992) reported that 65% of participants chose the one box response, and in a sample of 662 participants, Stanovich (1999) reported that 62.8% of participants chose the one box response. The findings here replicate the results reported by Shafir (1994) and Stanovich (1999; Stanovich & West, 2000), as a slightly larger proportion of the participants selected the Box B selection. While there has been some dispute about the optimal choice in this problem (Nozick, 1993), the two box consequentialist choice was taken as the disjunctive, optimal option, based on the dominance principle [Lewis, 1979; see also Stanovich’s (1999; Stanovich & West, 2000) demonstration of the understanding/acceptance principle in relation to this problem].

**Disease Framing Problem.** Originally used by Tversky and Kahneman (1981), the response pattern of no framing effect was taken as an index of disjunctive thinking, and 75.6% (n=93) of participants gave this response. Then, 21.1% (n=26) of participants displayed a framing effect. A framing effect is the selection of a risk averse choice in response to the prospect of losses, but a risk taking choice in response to the prospect of gains. Finally, 3.3% (n=4) of participants displayed a reverse framing effect. Reverse framers were not included in further analyses (see Methods for rationale).

Stanovich and West (1998a) reported that 69.2% (n=292) of their
participants displayed no framing effect in a within-participant administration of the Disease Framing Problem. Frisch (1993) found that 63.6% (n=73) of participants displayed no framing effect in the same within-participant administration of the Disease Framing Problem. The results of the current investigation indicate a higher proportion of no framing effect responders, but consistent in terms of the relative proportions of those who did or did not display a framing effect. Those who displayed the framing effect were treated as the disjunctive responders, and those who displayed no framing effect were treated as the non-disjunctive responders.

**Box Problem.** In this probabilistic judgment problem, 24% (n=30) of participants selected the 'no preference' option. Then, 67% (n=84) of participants selected the Game 1 response, and 9% chose the Game 2 response.

In Shafir's (1994) version of the Box Problem, participants were asked to choose between Game 1 and Game 2. Shafir found that 70% of participants selected Game 1 over Game 2, which is consistent with the current findings that the majority of participants selected Game 1. The preference for Game 1 was described by Shafir as evidence for violation of the disjunctive logic of the Sure Thing Principle. Although both Game 1 and Game 2 are equally likely, Shafir (1994) suggested that participants prefer to gamble on a simple event (Game 1) over an equally likely disjunctive event (Game 2).

Compared to the Box Problem used by Shafir, the version of this problem used in this study included an additional selection called the 'no preference' choice. The no preference response was treated as the disjunctive response in the current study, as this response reflects a recognition of the equivalence between the two games, and consequently, indifference between the two games. The Game 1 response was still treated as the non-disjunctive option,
Disjunctive Thinking

and Shafir (1994) also viewed this selection as the non-optimal option. The group of 11 participants who selected Game 2 in the current study were no longer the purely disjunctive responders in this study with the addition of the no preference response option. As Game 2 choosers demonstrate a preference for a more complex event over a simple event, they were given further consideration in some analyses because of trends that will be discussed later.

Informal Decision-Making Task

The Consumer Task. Participants generated an average of 5.4 (SD=2.32) features for the bread maker, and the total number of features generated ranged between 2 to 12 features. Based on Consumer Reports magazine, there were three diagnostic features which discriminated between different bread makers. Participants generated an average of 1.28 (SD=0.73) features which matched those reported by Consumer Reports. An average of 5.29 (SD=1.87) features were generated for the t-shirt, and the total number of features generated ranged between 1 to 12 features. Based on Consumer Reports magazine, there were four diagnostic features identified which discriminated different t-shirts, and participants generated an average of 1.58 (SD=0.79) features from Consumer Reports. An average of 8.04 (SD=2.86) features were generated for the car, and the number of features ranged between 2 to 18 features. Based on Consumer Reports magazine, there were four diagnostic features identified which discriminated different cars, and participants generated an average of 0.97 (SD=0.81) features from Consumer Reports.

These data are summarized in Table 5.

Figure 4 displays the frequency distribution of the number of unique features generated for each item. These distributions show that there was good variability in the number of features generated for each item, and the variability allows for fair comparisons of low generators with high
generators. A greater number of total features generated was used as a measure of disjunctive consideration.

**Formal Reasoning Problems**

**The Selection Task.** Table 6 displays the frequency of response selections on the selection task in the current study. Table 6 also displays frequencies from Stanovich and West (1998a; Stanovich, 1999) in a non-deontic with content version of the selection task called The Destination Problem. The results reported by Stanovich and West (1998a) overlap well with the distribution of responses found in this study. In this study, a smaller proportion of participants selected the \(P,Q\) matching response (40% vs. 49%), a smaller proportion of participants selected Other incorrect choices (13.6% vs. 18.0%), and a larger proportion of participants selected All of the cards (22.4% vs. 7.1%). No one selected \(P,Q,\text{Not}-Q\) in the sample reported here, but 4.8% of participants selected these cards in the sample reported by Stanovich and West (1998a). The \(P,\text{Not}-Q\) correct response was selected by only 7.2% of participants in this study. This is consistent with the original research done with the selection task, as the correct combination of \(P\) and \(\text{Not}-Q\) was reported to occur less than 10% of the time by Wason and Johnson-Laird (1972).

---

**Table 5**

**Mean Number of Features Generated in The Consumer Task**

<table>
<thead>
<tr>
<th>Item</th>
<th>Total Number of Features Generated</th>
<th>Features Generated from Consumer Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bread maker</td>
<td>5.40 (2.32)</td>
<td>1.28 (0.73)</td>
</tr>
<tr>
<td>T-Shirt</td>
<td>5.29 (1.87)</td>
<td>1.58 (0.79)</td>
</tr>
<tr>
<td>Car</td>
<td>8.04 (2.86)</td>
<td>0.97 (0.81)</td>
</tr>
</tbody>
</table>
Figure 4. Frequency Distribution of the Number of Unique Features Generated on The Consumer Task

Disjunctive Thinking
Table 6

Selection Task Performance Frequency Distribution in This Study and from Stanovich (1999; Stanovich & West, 1998)

<table>
<thead>
<tr>
<th>Response Selection</th>
<th>Current Study</th>
<th>Stanovich (1999; Stanovich &amp; West, 1998b) Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency Count (N=125)</td>
<td>Frequency Percent</td>
</tr>
<tr>
<td>Correct (P, Not-Q)</td>
<td>9</td>
<td>7.2%</td>
</tr>
<tr>
<td>P</td>
<td>21</td>
<td>16.8%</td>
</tr>
<tr>
<td>All</td>
<td>28</td>
<td>22.4%</td>
</tr>
<tr>
<td>P,Q</td>
<td>50</td>
<td>40.0%</td>
</tr>
<tr>
<td>P,Q, Not-Q</td>
<td>0</td>
<td>--</td>
</tr>
<tr>
<td>Other</td>
<td>17</td>
<td>13.6%</td>
</tr>
</tbody>
</table>

The extensive literature on the Selection Task reports some mixed findings on the optimal response selections in this task. For example, Margolis (1987) argued that choosing only the P card is an optimal selection. Stanovich (1999; Stanovich & West, 1998a) argued that the P, Not-Q response characterizes those who have experienced analytic success, and P,Q Choosers characterize heuristic responders. According to Stanovich (1999), P only responders, All card choosers, and P, Q, Not-Q responders characterize those who have experienced analytic failure. The arguments presented by Margolis and Stanovich provide a basis for examining P and All Choosers as alternative, optimal selections. The work of Margolis (1987) and Stanovich (1999) formed the rationale to collapse these different responses into one group of optimal, disjunctive responders. Therefore, those who provided responses reflecting analytic success or analytic failure were collapsed into one group
Disjunctive Thinking

(\text{Cor+P+All}), \text{and those who used the heuristic, matching strategy (P, Q) were}
the comparison group in the statistical analyses. The other incorrect
responders were eliminated from further analyses because it was unclear
whether their responses were disjunctive or non-disjunctive.

\textbf{Knights and Knaves Problem.} The relevant response options in this
problem were ‘knave’ and ‘cannot be determined.’ Of our respondents, 41.6%
(n=52) selected the knave response, and 48.8% (n=61) selected the cannot be
determined response. The choice of knave was treated as the disjunctive
solution to this problem, as consideration of all of the alternatives would result
in realizing that person C must be a knave. The cannot be determined
response choice was taken as the non-disjunctive response alternative.

Rips (1989) presented 34 participants with different Knights and Knaves
Problems, varying in difficulty (i.e., number of inference steps, where number
of inference steps increases with more speakers and more clauses). Rips
reported that performance ranged from 0% correct for the most difficult and
35% for the easiest version. Although Rips did not publish the actual problems
he used, the version used in the current investigation was likely the easiest,
least complex version that was used by Rips. If this is the case, 41.6% of
participants who correctly selected the knave response in the current study is
slightly higher, but similar to Rips’ findings with the easiest version of this
problem.

\textbf{Charles in Scotland Problem.} The frequencies of response selections are
presented in Table 7.

The Charles in Scotland problem has been studied as an instance of
propositional logic, and predictions about performance on this task have been
typically based on a mental models approach (Johnson-Laird & Byrne, 1991;
Johnson-Laird, Byrne, & Schaeken, 1992; Baron, 1994). Johnson-Laird et al.
Table 7

Frequencies of Response Selections on the Charles in Scotland Problem

<table>
<thead>
<tr>
<th>Response Selection</th>
<th>Frequency Count (N=125)</th>
<th>Frequency Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>16</td>
<td>12.8%</td>
</tr>
<tr>
<td>June/Kate</td>
<td>15</td>
<td>12.0%</td>
</tr>
<tr>
<td>Partial Contingency</td>
<td>15</td>
<td>12.0%</td>
</tr>
<tr>
<td>Nothing</td>
<td>11</td>
<td>8.8%</td>
</tr>
<tr>
<td>Blank</td>
<td>11</td>
<td>8.8%</td>
</tr>
<tr>
<td>Charles in Scotland</td>
<td>27</td>
<td>21.6%</td>
</tr>
<tr>
<td>Other Incorrect</td>
<td>30</td>
<td>24.0%</td>
</tr>
</tbody>
</table>

(1992) called the version used in this study as the “exclusive affirmative.” In their study, they predicted that this was the easiest version because there were only two possible models for each of the two premises, and that there were only two final models. The number of possible models for this problem is substantially constrained by the “but not both” phrasing at the end of each premise. In their results, Johnson-Laird et al. (1992) found that 21% of their participants solved this problem fully disjunctively, which is somewhat higher than the 12.8% who solved it fully disjunctively in this study.

The ‘nothing,’ blank responses, ‘Charles in Scotland,’ and other incorrect responses were less disjunctive in nature. However, the partial contingency and June/Kate responders were not clearly disjunctive or non-disjunctive, and were therefore given further consideration. In order to arrive at a partial contingency response, this solution requires temporarily assigning truth values to each disjunct in each premise. That is, Charles CAN be in Scotland but then June CANNOT be in Wales — this is the very step
that Shafir (1994) identifies as the foregone step in this problem. Shafir (1994) states that “presented with a disjunction of simple alternatives most subjects refrain from assuming the respective disjuncts and arrive at no valid conclusions” (p. 423). A partial contingency response therefore involves some disjunctive consideration, or navigation down at least one branch of the decision tree. In order to provide a partial contingency response, a virtually exhaustive consideration of the two possible models would have taken place (Johnson-Laird, Byrne, & Schaeken, 1992). A June/Kate response, a less high level response, reflects some amalgamation across the two mental models (Johnson-Laird et al., 1992). That is, ‘June is in Wales’ and ‘Kate is in Ireland’ are separate premises, and this solution reflects some consideration of the two possible models and assignment of the same truth value to each of these premises. Therefore, some of the analyses collapsed the partially contingent and June/Kate with the fully disjunctive, Correct responders.

Therefore, three sets of comparisons were used in the statistical analyses of the Charles in Scotland Problem in order to explore trends across the different responses. First, based on a priori comparisons, Correct responders and Partial Contingency responders were collapsed (Cor+Partial) to form a highly analytic group and compared with all the undifferentiated other responses (June/Kate+Charles+Nothing+Blank+Other). Second, the correct, partial contingency, and June/Kate responders were collapsed into a broadly analytic group and compared with all the other responses, called the differentiated other group (Charles+Nothing+Blank+Other). Third, the correct, partial contingency, and June/Kate (Cor+Partial+June/Kate) responders were collapsed and compared with the Charles is in Scotland responders. The third analysis was the sharpest comparison because the broadly analytic (Cor+Partial+June/Kate) responders were compared with the
best approximation of an heuristic response: the Charles in Scotland (Charles) responders. In summary, the following comparisons were examined:

1. ‘Highly Analytic’ versus ‘Undifferentiated Other’ 
   (Cor+Partial) versus (June/Kate+Charles+Nothing+Blank+Other); 

2. ‘Broadly Analytic’ versus ‘Differentiated Other’ 
   (Cor+Partial+June/Kate) versus (Charles+Nothing+Blank+Other); 
and

3. ‘Broadly Analytic’ versus ‘Heuristic’ 
   (Cor+Partial+June/Kate) versus (Charles).

These different comparisons provide a richer and more in-depth analysis of the disjunctive aspects of the Charles in Scotland Problem. If partial contingency and June/Kate responders display similar tendencies as those who provided a fully disjunctive response, then this would provide a richer conceptualization of disjunctive thinking, which includes different gradients or levels of disjunctive consideration.

**Married Problem.** The relevant response alternatives in this problem were ‘yes’ and ‘cannot be determined.’ Only 12.8% (n=16) participants selected the yes response, and 85.6% (n=107) selected the cannot be determined response. The response selection yes was treated as the disjunctive choice, because a true yes response should signal those participants who realized that whether Ann was married or unmarried, either way, a married person would be looking at an unmarried person. The cannot be determined alternative was treated as the non-disjunctive response. That is, those participants who failed to travel all the branches of the tree -- not realizing that the solution was yes if Ann was married, and yes if Ann was unmarried.

What is striking about this problem is the low frequency of participants who selected the disjunctive yes response. Relative to performance on the other reasoning and decision making problems (see Table 4), performance was
Disjunctive Thinking

lowest on the Married Problem and the Green Levels Problem. Both of these problems were taken from Levesque (1986, 1989). Levesque's (1986) discussion of these problems was not in terms of human performance, but rather in terms of computer modeling and simulation. Levesque did argue that the conclusion to be drawn in this problem is implicitly stated. It is Ann's marital status that is implicit in this problem, and it is this step which may be difficult to realize in this problem. The options for Ann's status can only be 'married' or 'unmarried', as these options exhaust the possible alternatives. These parameters suggest that it is a typical deductive logic problem, but it is a puzzle as to why so few participants selected the disjunctive response.

**Green Levels Problem.** The relevant response alternatives were 'yes' and 'cannot be determined.' Of our 125 participants, 8.8% (n=11) selected the yes response, and 84% (n=105) of participants selected the cannot be determined response alternative. A yes response was interpreted as the disjunctive response, and the cannot be determined response choice was treated as the non-disjunctive alternative.

The logic of the Green Levels Problem is analogous to the Married Problem. Similar to the Married Problem, Levesque (1989) introduced this logic problem in the context of computer simulation. As compared to the Married Problem, more inference and reasoning are needed in this problem, as there are three unknowns that must be inferred: the top, middle, and bottom blocks. The Green Problem did seem to be more difficult, as only 8.8% of participants responded correctly on the Green Levels problem as compared to 12.8% of participants on the Married Problem. Both problems are characterized by a very low frequency of disjunctive responders.

**Informal Reasoning Task**

**Argument Generation Task.** Participants generated a mean of 5.65
Disjunctive arguments for the tuition issue, 4.90 (SD=1.67) arguments for the organs issue, and 4.96 (SD=1.74) arguments for the gasoline issue. Of interest in the current study was the number of arguments which endorsed participants' prior belief (my-side arguments) and the number of arguments which did not endorse participants' prior belief (other-side arguments) on the issues. An equivalent number of my-side and other-side arguments or a higher proportion of other-side arguments were taken as a reflection of disjunctive consideration. The mean number of my-side and other-side arguments for each issue is presented in Table 8. For the tuition issue, there were significantly more my-side than other-side arguments generated \[t(127)=8.5, p<.0001\]. Similarly for the organ issue, there were significantly more my-side than other-side arguments generated \[t(127)=3.7, p<.0001\]. Also for the gasoline issue, there were significantly more my-side than other-side arguments generated \[t(127)=4.19, p<.0001\].

Table 8 also includes the average belief bias index for each issue. The belief bias index is the discrepancy score between the number of my-side and other-side arguments for each issue. The belief bias index was significantly greater than zero for all three issues, demonstrating the presence of belief bias

<table>
<thead>
<tr>
<th>Issue</th>
<th>&quot;My-side&quot; Generated</th>
<th>&quot;Other side&quot; Generated</th>
<th>Belief Bias Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuition</td>
<td>3.38 (1.30)</td>
<td>2.14 (1.02)</td>
<td>1.24 (1.65)</td>
</tr>
<tr>
<td>Organs</td>
<td>2.67 (1.16)</td>
<td>2.13 (1.14)</td>
<td>0.54 (1.64)</td>
</tr>
<tr>
<td>Gasoline</td>
<td>2.74 (1.24)</td>
<td>2.15 (1.07)</td>
<td>0.59 (1.60)</td>
</tr>
</tbody>
</table>
on the tuition issue \( \chi^2(1) = 8.5, p < .0001 \), on the organ issue \( \chi^2(1) = 3.72, p < .0001 \), and on the gasoline issue \( \chi^2(1) = 4.19, p < .0001 \). Positive scores on the belief bias index indicate that more my-side than other-side arguments were generated for each issue. The distribution of the belief bias scores for each of the three issues is shown in Figure 5.

The results in Figure 5 illustrate that there was an overall tendency to not reason disjunctively, as positive belief bias scores indicate that more my-side than other-side arguments were generated.

2. Domain Generality

Formal Decision-Making and Formal Reasoning Problems

A series of chi-square analyses were performed to examine response tendencies across different reasoning problems in order to examine the domain generality hypothesis. It was of interest to determine whether a disjunctive response on one problem was statistically associated with a disjunctive response on another problem. Table 9 displays the phi coefficients obtained between all of the formal decision-making and reasoning problems. Phi coefficients are a special case of the product-moment correlation \( r \) for dichotomous variables, and they can be interpreted as correlations (Rosenthal & Rosnow, 1991). Phi coefficients are obtained from chi-square analyses.

Knights and Knaves, Married Problem, Green Levels Problem. The chi-square analyses displayed a nice triangulation of response tendencies in the Knights and Knaves Problem, the Married Problem, and the Green Levels Problem. Disjunctive response selections on each of these problems were associated with disjunctive response selections on the other problems. That is, ‘knave’ responders on the Knights and Knaves Problem were more likely to respond ‘yes’ on the Married Problem, and ‘cannot be determined’ responders in each problem were also positively associated \( \chi^2(1) = 9.38, p < .01 \). The same
Figure 5. Frequency Distribution of Belief Bias Indices on The Argument Generation Task
Table 9

<table>
<thead>
<tr>
<th>Variable</th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decision-Making Tasks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Prisoner’s Dilemma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Newcomb’s Problem</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Framing Problem</td>
<td>.05</td>
<td>.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Box Problem</td>
<td>.10</td>
<td>.16</td>
<td>.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Reasoning Tasks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Selection Task</td>
<td>.05</td>
<td>.00</td>
<td>.09</td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Knights and Knaves</td>
<td>.06</td>
<td>.10</td>
<td>.07</td>
<td>.07</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Charles in Scotland</td>
<td>.08</td>
<td>.08</td>
<td>.25*</td>
<td>.28*</td>
<td>.31**</td>
<td>.24*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Married Problem</td>
<td>.11</td>
<td>.04</td>
<td>.09</td>
<td>.09</td>
<td>.00</td>
<td>.29*</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>9. Green Problem</td>
<td>.00</td>
<td>.06</td>
<td>.03</td>
<td>.10</td>
<td>.05</td>
<td>.19†</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.31**</td>
</tr>
</tbody>
</table>

† Coefficients larger than .176 are significant at the .05 level (two-tailed);
* Coefficients larger than .230 are significant at the .01 level (two-tailed); and
** Coefficients larger than .291 are significant at the .001 level (two-tailed)

Note: The following comparisons were made, and the disjunctive response is italicized.
Prisoner’s Dilemma: ‘Compete’ vs. ‘Cooperate’
Newcomb: ‘Both Boxes’ vs. ‘Box B Only’
Framing Problem: ‘No Framing Effect’ vs. ‘Framing Effect’
Box Problem: ‘No Preference’ vs. ‘Game 1’
Selection Task: ‘Cor+P+All’ vs. ‘PQ’
Knights and Knaves: ‘Knave’ vs. ‘CBD’
Charles in Scotland: ‘Cor+Partial+June/Kate’ vs. ‘Charles in Scotland’
Married Problem: ‘Yes’ vs. ‘CBD’
Green Levels Problem: ‘Yes’ vs. ‘CBD’.
result was obtained between the Knights and Knaves Problem and the Green Levels Problem. That is, 'knave' responders were more likely to choose the disjunctive response on the Green Levels problem, and 'cannot be determined' responders were positively associated on both problems \( \chi^2(1) = 3.7, p < .05 \).

Finally, 'yes' responders were highly associated on the Married Problem and the Green Levels Problem, as were 'cannot be determined' responders \( \chi^2(1) = 11.265, p < .001 \). The statistical relationships obtained suggest that performance on these three problems are related. Based on a priori discussions, they share some aspect of disjunctive consideration. In addition, this particular pattern of associations between the reasoning problems also suggests that there is a problem solving element embedded in these problems which importantly ties them together.

**Charles in Scotland Problem.** Response selections on the Charles in Scotland Problem were significantly associated with response choices on two other formal reasoning problems: the Selection Task and the Knights and Knaves Problem. Also, responses on the Charles in Scotland Problem were significantly associated with response choices on two formal decision-making problems: the Box Problem, and the Framing Problem. The broadly analytic responders (Correct+Partial+June/Kate) and the heuristic responders (Charles in Scotland) in the Charles in Scotland Problem were included in these analyses. Some analyses were performed with the other response groupings to examine differences in trends between the different groupings.

Response tendency relationships between the Charles in Scotland and the other formal reasoning problems were in the anticipated direction. Specifically, those who gave a broadly analytic response on the Charles in Scotland Problem were more likely to select a more disjunctively oriented
response (P+Cor+All) on the Selection Task [$\chi^2(1) = 6.12, p<.01$]. Also, those who provided a heuristic response on the Charles in Scotland Problem were more likely to select the heuristic P,Q response on the Selection Task. The broadly analytic responders on the Charles in Scotland Problem were more likely to select the ‘knave’ response on the Knights and Knaves Problem [$\chi^2(1) = 3.86, p<.05$]. Then, heuristic responders in the Charles in Scotland Problem were more likely to select the ‘cannot be determined’ response on the Knights and Knaves Problem.

Similarly, response tendencies on the Charles in Scotland Problem and the formal decision-making problems were also in the anticipated direction. Those who gave a broadly analytic response on the Charles in Scotland Problem were more likely to choose the ‘no preference’ choice in the Box Problem [$\chi^2(2) = 6.2, p<.05$]. Also, Charles in Scotland responders were more likely to select Game 1 or Game 2 in the Box Problem. Those who gave the broadly analytic response on the Charles in Scotland Problem were less likely to display the framing effect in the Disease Framing Problem [$\chi^2(1) = 4.27, p<.05$]. Conversely, Charles in Scotland responders were more likely to display the framing effect.

Other Comparisons With The Charles in Scotland Problem. When the analyses were done with the other partitioning of groups in the Charles in Scotland Problem, some of these relationships were obtained (please see page 69 for the different grouping of responses in the Charles in Scotland Problem). In the broadly analytic versus undifferentiated other and the broadly analytic versus differentiated other comparisons, no significant chi-square relationships were obtained with the Selection Task. In the broadly analytic versus differentiated other comparison, the broadly analytic group in the
Charles in Scotland Problem displayed a marginally significant tendency to select 'knave' in the Knights and Knaves Problem [$\chi^2(1) = 2.74, p<.10$]. In the same analysis, the differentiated other group tended to choose 'cannot be determined' on the Knights and Knaves Problem. No differences were obtained with the Knights and Knaves Problem when the comparison included the undifferentiated other group in the Charles in Scotland Problem.

On the Box Problem, both sets of chi-square comparisons were significant in the expected direction, with the undifferentiated other comparison [$\chi^2(1) = 7.27, p<.05$], and the differentiated other comparison [$\chi^2(1) = 6.88, p<.05$]. That is, 'no preference' choosers on the Box Problem were more likely to choose a broadly analytic response on the Charles in Scotland Problem, and Game 1 and Game 2 choosers were more likely to select an undifferentiated other or differentiated other response.

The broadly analytic group in the Charles in Scotland Problem had a marginally significant tendency to display no framing effect in the Disease Framing Problem [$\chi^2(1) = 2.76, p<.10$]. In the same analysis, the differentiated other group tended to display a framing effect in the Disease Framing Problem. No differences were obtained with the Disease Framing Problem when the comparison included the undifferentiated other group in the Charles in Scotland Problem. These additional comparisons suggest that the allocation of the June/Kate, nothing, blank, and other incorrect responders in the Charles in Scotland Problem is very important as they seem to affect the pattern of results considerably. The comparison of the highly analytic versus the heuristic responders in the Charles in Scotland Problem with the other formal problems displayed the most frequent significant results.
Selection Task. Response choices on the Selection Task were significantly related to responses in the Charles in Scotland Problem in the expected direction. Some additional analyses were performed with P only choosers and All choosers on the Selection Task. As there was a sizable group of P only choosers and All choosers, separate chi-square analyses were performed. These analyses were done to further investigate whether P only choosers and All choosers are disjunctive response selections.

Selection Task: P Only Choosers versus P,Q Choosers. In a separate analysis, P choosers were compared with P,Q choosers in a set of chi-square analyses. The results of these analyses and the phi coefficients are displayed in Table 11. A significant relationship emerged with the Framing Problem, and marginal relationships emerged with the Knights and Knaves Problem and the Prisoner’s Dilemma. In the Framing Problem, P only choosers were less likely to show a framing effect and P,Q choosers were more likely to display a framing effect \( \chi^2(1)=5.002, p<.025 \). On the Knights and Knaves problem, P only choosers were more likely to select the correct knave response, while P,Q choosers were more likely to select the heuristic ‘cannot be determined’ response alternative \( \chi^2(1)=3.03, .05<p<.10 \), although this effect was marginal. On the Prisoner’s Dilemma, P only choosers were more likely to select the compete disjunctive solution, while P,Q choosers were more likely to select the cooperate less disjunctive solution, but this effect was marginal \( \chi^2(1)=2.84, .05<p<.10 \). What is notable here is that P only choosers tended to provide a disjunctive response on other formal reasoning and decision-making problems. This is consistent with theoretical discussions (Margolis, 1987) and other empirical demonstrations (Stanovich, 1999; Stanovich & West, 1998a) of the Selection Task.
Selection Task: All Choosers versus P,Q Choosers. No relationships were obtained with All choosers, as compared with the relationships that were obtained with P only choosers. The results of these analyses and the phi coefficients are displayed in Table 10. The results presented in Table 10 suggest that All choosers reflect a less disjunctive response than displayed by P only choosers.

Table 10

Phi Coefficients with P Only and All Choosers in the Selection Task

<table>
<thead>
<tr>
<th></th>
<th>P Only Choosers versus P,Q Choosers</th>
<th>All Choosers versus P,Q Choosers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Formal Decision-Making Problems</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prisoner's Dilemma</td>
<td>0.20†</td>
<td>0.03</td>
</tr>
<tr>
<td>Newcomb's Problem</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>Disease Framing Problem</td>
<td>0.27*</td>
<td>0.02</td>
</tr>
<tr>
<td>Box Problem</td>
<td>0.21</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Formal Reasoning Problems</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knights &amp; Knaves</td>
<td>0.22†</td>
<td>0.10</td>
</tr>
<tr>
<td>Charles in Scotland</td>
<td>0.33*</td>
<td>0.20</td>
</tr>
<tr>
<td>Married Problem</td>
<td>0.15</td>
<td>0.08</td>
</tr>
<tr>
<td>Green Problem</td>
<td>0.14</td>
<td>0.02</td>
</tr>
</tbody>
</table>

† = p < .10, * = p < .05
Box Problem, Disease Framing Problem, Newcomb, Prisoner's Dilemma. Shafir (1994) did not include the Disease Framing Problem in his discussion of disjunctive thinking, however, it was expected that this framing problem would fit into the broad category of tasks involving disjunctive consideration. In particular, it was expected that the Disease Framing Problem required a similar type of reversibility in thinking as the Box Problem. Namely, if one calculates the expected value of both alternatives in the Framing Problem and the Box Problem, participants should realize the relative equivalence of the options in both problems. No such relationship was found. Both the Framing Problem and the Box Problem seem to necessitate disjunctive consideration, but the evidence did not support this expected relationship.

Newcomb's Problem and the Prisoner's Dilemma, in addition to Shafir (1994), have been linked in the philosophical literatures. While both problems involve disjunctive consideration, these two problems have also been also necessitate a type of 'magical thinking' (Shafir, 1994). That is, participants may attribute a selected response to the unknown 'Other', which in Newcomb's Problem would be the Computer Program, and in the Prisoner's Dilemma, it would be the other game player. An instantiation of magical thinking may lead one to assume that the Computer Program's prediction is correct in Newcomb's Problem, and that the other player will cooperate in the Prisoner's Dilemma. However, the disjunctive response selection would be to compete in the Prisoner's Dilemma, and the selection of both boxes in the Newcomb problem. These choices are the optimal solutions because they would serve to maximize the total expected value. No such relationships were found between the response selections on these two problems in the chi-square analyses.
In light of Shafir's discussion, response selections across the Box Problem, the Disease Framing Problem, Newcomb's Problem, and the Prisoner's Dilemma were compared. One comparison approached significance. That is, response choices on the Disease Framing Problem were significantly related to response choices on Newcomb's Problem $\chi^2(1) = 3.4$, $p<.06$. Namely, those who selected the one box solution in Newcomb's Problem were more likely to display a framing effect, and those who selected both boxes were more likely to display no framing effect.

Together, these findings display some of the expected patterns: there is some evidence for domain generality within the formal reasoning problems, and only some links between the broad domains of the formal reasoning and decision-making problems. In particular, relationships between response tendencies were obtained on the Knights and Knaves Problem, Married Problem, Green Levels Problem, Charles in Scotland, and the Selection Task (formal reasoning problems), and between Charles in Scotland, Box Problem, and Disease Framing Problem (formal reasoning and decision-making problems). Some of the expected relationships were not obtained, for example, between the Prisoner's Dilemma and Newcomb's Problem, which suggest that there may be other important aspects involved in solving these problems. While the alternative courses of action must be considered and evaluated in each of these problems, disjunctive consideration itself may not be enough to reach the optimal solution on these problems. There is certainly a history in the philosophical literature to support this suggestion (Nozick, 1993).

Informal Tasks

Consumer Task. The total number of features generated for the bread maker, t-shirt, and car were subjected to correlational analyses in order to examine domain generality within this task. That is, did those who generate
more features for one item display a tendency to generate more features for the other items? These results are displayed in Table 11.

Table 11

Domain Generality within Consumer Task: Correlations Between Number of Features Generated

<table>
<thead>
<tr>
<th></th>
<th>Bread maker</th>
<th>T-Shirt</th>
<th>Car</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bread maker</td>
<td>--</td>
<td>0.55***</td>
<td>--</td>
</tr>
<tr>
<td>T-Shirt</td>
<td>0.55***</td>
<td>--</td>
<td>0.52***</td>
</tr>
<tr>
<td>Car</td>
<td>0.52***</td>
<td>0.52***</td>
<td>--</td>
</tr>
</tbody>
</table>

*** = p < .001

As shown in Table 11, a significant correlation was found between the number of unique features generated for the bread maker and t-shirt (r=.545, p<.001), between the number of unique features generated for the bread maker and car (r=.515, p<.001), and between the number of unique features generated for the t-shirt and car (r=.522, p<.001). These correlations indicate that those who generated more features for one of these items was more likely to generate more features for the other items. Consequently, these correlations suggest that there is domain generality within this task, with three very different items. Given these associations, a composite with the total number of unique features for all three items was constructed for further statistical analyses (Consumer Task - Total Unique Features).

Argument Generation Task. Domain generality in the Argument Generation Task was explored in three different ways. First, analogous to the Consumer Task, the total number of arguments generated for each issue was subjected to correlational analyses. Second, the number of my-side and other-
side arguments for each issue were correlated. Finally, the belief bias index constructed for each issue was correlated across the different issues. In a simple correlational analysis, it was of interest to examine whether undifferentiated argument generation was related across the three different issues. These correlations are displayed in Table 12.

Table 12
The Argument Generation Task: Correlations Between Total Number of Arguments Generated For Each Issue

<table>
<thead>
<tr>
<th></th>
<th>Tuition</th>
<th>Organs</th>
<th>Gasoline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuition</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organs</td>
<td>0.58***</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Gasoline</td>
<td>0.63***</td>
<td>0.61***</td>
<td>-</td>
</tr>
</tbody>
</table>

*** = p < .001

The correlations in Table 12 indicate that participants who generated more arguments for one issue had a tendency to generate more arguments for the other issues. These correlations suggest that this generational tendency spans across all three issues, supporting a type of domain generality that is related to argument generation.

Table 13 displays the correlations between the number of my-side and other-side arguments generated for each of the three issues. There are two main sets of findings that are of interest in Table 13. First, the numbers of my-side arguments are positively related between: the tuition and organ issues ($r=.342, p<.01$), the tuition and gasoline issues ($r=.411, p<.01$), and the organ and gasoline issues ($r=.343, p<.01$). The numbers of other-side arguments are positively related for the tuition and organ issues ($r=.357, p<.01$) and for the tuition and gasoline issues ($r=.395, p<.01$). There are also some significant
### Domain Generality within Argument Generation Task: Correlations Between Number of My-side and Other-side Arguments Generated

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Tuition My-side</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Tuition Other-side</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Organ My-side</td>
<td>0.32***</td>
<td>0.36***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Organ Other-side</td>
<td>0.19*</td>
<td>0.35***</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Gas My-side</td>
<td>0.37***</td>
<td>0.27**</td>
<td>0.36***</td>
<td>0.43***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Gas Other-side</td>
<td>0.24**</td>
<td>0.38***</td>
<td>0.36***</td>
<td>0.10</td>
<td>-0.10</td>
<td></td>
</tr>
</tbody>
</table>

* = p < .05, ** = p < .01, *** = p < .001

Correlations between the number of my-side and other-side arguments across the different issues: (1) the number of my-side arguments generated for the organs issue was significantly related to the number of other-side arguments generated for the tuition issue ($r = .292, p < .01$), (2) the number of my-side arguments generated for the gasoline issue was significantly related to the number of other-side arguments for the tuition issue ($r = .292, p < .05$), (3) and the number of my-side arguments for the gasoline issue was significantly related to the number of other-side arguments generated for the organs issue ($r = .43, p < .01$). Overall, these correlations demonstrate some domain generality of argument generation fluency, which is consistent with the
previously reported correlations between total arguments generated for each issue. That is, participants who generated more arguments for one issue (either my-side or other-side) were more likely to generate more arguments (either my-side or other-side) for other issues. Given these associations, composites with the total number of arguments (Argument Generation Task - Total Arguments), total number of my-side arguments (Argument Generation Task - Total my-side Arguments), and total number of other-side arguments (Argument Generation Task - Total other-side Arguments) for all three issues were constructed for further statistical analyses.

The second interesting finding is that there was no significant relationship between the number of my-side and other-side arguments within any one issue (see Table 14). This finding suggests that despite the hint of domain generality in generating arguments across issues, there was no evidence of domain generality in generating my-side and other-side arguments within an issue. That is, those who generated more my-side arguments for the tuition issue may have generated more other-side arguments for the organ issue, but did not generate more other-side arguments for the tuition issue. Therefore, there is evidence to support generality akin to some type of generational fluency both within and across the different issues, but not when it comes to weighing the arguments equally for a given issue. Interesting, this finding leads to a more in depth conceptualization of disjunctive thinking. That is, disjunctive consideration is necessary for the generation of arguments, but as conceptualized in the present study, disjunctive thinking is also important for the equal weighting of arguments in a given issue. In particular, this finding has interesting implications for conceptualizing disjunctive reasoning in informal tasks.

Importantly, this tendency to generate was observed on all three issues, and
the lack of relationships between my-side and other-side arguments was also observed across all three issues used in this task.

Finally, Table 14 displays correlations between the belief bias indices across the three different issues in the Argument Generation Task.

Table 14
Correlations Between Belief Bias Indices on The Argument Generation Task

<table>
<thead>
<tr>
<th></th>
<th>Tuition</th>
<th>Organs</th>
<th>Gasoline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuition</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organs</td>
<td>0.07</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Gasoline</td>
<td>0.13</td>
<td>-0.16</td>
<td>--</td>
</tr>
</tbody>
</table>

As displayed in Table 14, there were no significant associations between the belief bias indices on the three issues. Significant correlations would have indicated that those who displayed a belief bias on one issue displayed belief bias tendencies on the other two issues. Alternatively, these findings show a lack of generalizability in belief bias tendencies on this task. That is, those who displayed a strong belief bias on the tuition issue, as measured by more my-side than other-side arguments generated, are not necessarily those who displayed a strong belief bias on the organs issue or the gasoline issue. The lack of relations in Table 14 are not inconsistent with the generational tendencies displayed in Tables 12 and Table 13.

To summarize the results presented in Tables 12, 13, and 14, the correlations in Tables 12 and 13 suggest that those who tended to display more arguments for one issue had a tendency to generate more arguments for another issue. This generational tendency was also true for the number of my-
side and other-side arguments that were generated, but not within a given issue. However, the results in Table 14 suggest that the ratio of the number of my-side to other-side arguments within each issue are not significantly related. For example, while some participants may have provided a proportionately larger number of my-side arguments for the tuition issue, those same participants may not have provided such a large proportion of my-side arguments for the gasoline issue. While the results support some generality of belief bias, as evidenced by the larger number of my-side than other-side arguments generated overall (see Table 8), there is also some evidence for specificity across the three issues used here. There is evidence for my-sided-ness across the three issues, but the extent of mysided-ness seems to be different for each issue.

**Consumer Task and Argument Generation Task.** In order to explore patterns of response tendencies on the Consumer Task and the Argument Generation Task, correlational analyses were performed between the number of features generated for the Consumer Task and the total number of arguments generated in the Argument Generation Task. These results are displayed in Table 15.

The correlations in Table 15 display generational tendencies between the Consumer Task and the Argument Generation Task. The total number of features generated in the Consumer Task were significantly correlated with the total number of arguments generated for the Argument Generation Task ($r=0.48$, $p<.001$). All of the other relationships were also significant, again suggesting a type of generational fluency within and across these tasks - those who tend to generate more features or arguments for one item or issue display this tendency for other items and issues. The correlations in Table 15 provide further evidence for a form of generational fluency on the informal tasks. This
Correlations Between Features and Arguments Generated on The Consumer Task and The Argument Generation Task

<table>
<thead>
<tr>
<th></th>
<th>Total Features</th>
<th>Consumer Reports Features</th>
<th>Total Arguments</th>
<th>My-side Arguments</th>
<th>Other-side Arguments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Features</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumer Reports Features</td>
<td>0.61***</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Arguments</td>
<td>0.48***</td>
<td>0.33***</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>My-side Arguments</td>
<td>0.42***</td>
<td>0.26**</td>
<td>0.87***</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Other-side Arguments</td>
<td>0.32***</td>
<td>0.27**</td>
<td>0.80***</td>
<td>0.42***</td>
<td>--</td>
</tr>
</tbody>
</table>

** = p < .01, *** = p < .001

tendency has been displayed in The Consumer Task (Table 11), in The Argument Generation Task Table 12 and Table 13), and between The Consumer Task and The Argument Generation Task (Table 15).

** Formal and Informal Tasks**

The same statistical strategy was used in order to explore relationships between response tendencies on the formal decision-making/reasoning problems and the informal decision-making/reasoning tasks. For the Consumer Task, the total number of unique features generated and the total number of features matched from Consumer Reports were compared on the disjunctive and non-disjunctive response alternatives for each formal
decision-making and reasoning problem. For the Argument Generation Task, the total number of my-side arguments, the total number of other-side arguments, and the total belief bias were compared on the disjunctive and non-disjunctive response alternatives for each decision-making and reasoning problem.

The Consumer Task and Formal Decision-Making and Reasoning Problems. No relationships were found between any of the decision-making or reasoning problems and the total number of unique features or with the total number of Consumer Reports features generated. This suggests that disjunctive response selections on the formal decision-making and reasoning problems were not related to generation on the Consumer Task. This provides further evidence for the previously described discrimination between generation and evaluation: there are elements of both generation and evaluation involved in disjunctive thinking. However, simply generating the alternatives is not sufficient to derive the correct solution in the formal problems. For example, it is insufficient to realize that Ann could be married or unmarried in the Married Problem without considering the implications of each in the problem space. The data presented from the informal tasks have provided evidence for a generational tendency. The lack of relations between the decision-making/reasoning problems and the Consumer Task may be capturing this important distinction between generation and evaluation within the domain of disjunctive thinking.

The Argument Generation Task and Formal Decision-Making and Reasoning Problems. No relationships were obtained between any of the formal tasks and The Argument Generation Task. Relationships were examined with the total number of my-side arguments, total number of other-side arguments, and with total belief bias. This suggests that disjunctive
Disjunctive Thinking

response solutions on the formal problems are not directly related to argument generation in the Argument Generation Task. Further, disjunctive solutions on the formal problems were not related to the number of my-side or other-side arguments generated in the Argument Generation Task. The number of my-side and other-side arguments generated were intended to capture the evaluative notion of disjunctive thinking, that is, treating both sides of the issue equivalently.

These findings suggest that neither argument generation nor the evaluative (my-side and other-side) aspects of the Argument Generation Task were related to performance on the formal tasks. These results provide evidence for the lack of generality between the formal and informal tasks used in this study.

3. Predicting Performance on Formal and Informal Tasks using Cognitive Ability and Thinking Dispositions

A series of simple comparisons were performed for each formal decision-making and reasoning problem in order to explore patterns of relationships with cognitive ability and thinking dispositions. The measures of Total Cognitive Ability (Verbal Cognitive Ability + Nonverbal Cognitive Ability), Verbal Cognitive Ability, and Nonverbal Cognitive Ability were the measures of cognitive ability. The focal thinking disposition measures included the Actively Open Minded Thinking Composite (AOT), the Need for Cognition Composite (Need for Cognition Scaled Score + Typical Intellectual Engagement Scaled Score), the Need for Cognition Total Raw Score, and the Matching Familiar Figures Test (MFFT). Higher scores on the AOT and NeedCog measures are positively coded; for example, a higher score on the AOT scale indicates that participants endorsed items related to actively open-minded thinking. Three different metrics from the Matching Familiar
Disjunctive Thinking

Figures Test (MFFT) were used: total reaction time to first response (MFFT_{RT}), total number of errors (MFFT_{Errors}), and the discrepancy score between reaction time and errors (MFFT_{RT-Errors}). A positive, high score on the reaction time and discrepancy measures are indicators of a more reflective style, and negative scores on MFFT_{Errors} is indicative of a reflective style.

Means and standard deviations on all of the individual difference measures for each formal decision-making and reasoning problem are presented. Where relevant, significant effects on other thinking dispositions scales are reported in the body of the results. Table 16 presents a summary of the means and standard deviations on Total Cognitive Ability, MFFT_{RT-Errors}, and the Need for Cognition Total Raw Score. Table 17 provides a summary of the significant differences and trends that were obtained on all of the individual difference measures.

**Formal Decision-Making Problems**

The results of the simple comparisons with the cognitive ability and thinking dispositions measures are presented in a separate table for each formal decision-making and reasoning problem. Disjunctive and non-disjunctive responses, as described in the Method section, were compared on each individual difference measure.

**Prisoner's Dilemma.** Compete and cooperate responders were compared on the ability and dispositional measures, and these comparisons are presented in Table 18.

On the cognitive ability measures, those who chose to compete had a higher Nonverbal Cognitive Ability score than those who chose to cooperate. As compete responders are the disjunctive responders, this effect is in the anticipated direction. No other ability differences were found. There were no differences on the AOT and NeedCog scales, nor on the MFFT measures. The
### Table 16

**Total Cognitive Ability, MFFT<sub>RT</sub>-Errors<sup>-</sup> - and Need for Cognition Total Raw Score Means and Standard Deviations on Formal Tasks**

<table>
<thead>
<tr>
<th></th>
<th>Disjunctive</th>
<th>Non-Disjunctive</th>
<th>Disjunctive</th>
<th>Non-Disjunctive</th>
<th>Disjunctive</th>
<th>Non-Disjunctive</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Formal Decision-Making Problems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prisoner's Dilemma</td>
<td>0.31 (2.6)</td>
<td>-0.30 (3.0)</td>
<td>-0.07 (1.8)</td>
<td>0.11 (1.6)</td>
<td>77.5 (10.8)</td>
<td>79.5 (10.3)</td>
</tr>
<tr>
<td>Newcomb's Problem</td>
<td>-0.16 (2.6)</td>
<td>0.12 (2.9)</td>
<td>-0.01 (1.7)</td>
<td>0.01 (1.8)</td>
<td>77.4 (11.9)</td>
<td>78.66 (10.1)</td>
</tr>
<tr>
<td>Disease Problem</td>
<td>0.20 (3.0)</td>
<td>-0.56 (2.1)</td>
<td>-0.01 (1.8)</td>
<td>0.06 (1.7)</td>
<td>80.0 (10.3)</td>
<td>74.27 (10.0)&lt;sup&gt;**&lt;/sup&gt;</td>
</tr>
<tr>
<td>Box Problem</td>
<td>1.06 (3.3)</td>
<td>-0.22 (2.5)&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.42 (1.5)</td>
<td>0.02 (1.8)</td>
<td>78.6 (9.8)</td>
<td>76.76 (10.7)</td>
</tr>
<tr>
<td><strong>Formal Reasoning Problems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection Task</td>
<td>0.75 (2.7)</td>
<td>-0.46 (2.3)&lt;sup&gt;**&lt;/sup&gt;</td>
<td>0.50 (1.6)</td>
<td>-0.30 (1.9)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>79.76 (11.0)</td>
<td>76.98 (10.5)</td>
</tr>
<tr>
<td>Knights &amp; Knaves</td>
<td>0.57 (2.8)</td>
<td>-0.03 (2.7)</td>
<td>0.32 (1.6)</td>
<td>-0.18 (1.7)</td>
<td>81.25 (11.3)</td>
<td>76.69 (9.8)&lt;sup&gt;**&lt;/sup&gt;</td>
</tr>
<tr>
<td>Charles in Scotland</td>
<td>1.15 (2.5)</td>
<td>-1.53 (2.4)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.45 (1.6)</td>
<td>-0.71 (1.9)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>81.2 (11.2)</td>
<td>77.78 (8.4)</td>
</tr>
<tr>
<td>Married Problem</td>
<td>-0.61 (2.9)</td>
<td>0.14 (2.7)</td>
<td>-0.32 (2.0)</td>
<td>0.02 (1.7)</td>
<td>81.19 (10.1)</td>
<td>77.85 (10.9)</td>
</tr>
<tr>
<td>Green Problem</td>
<td>1.03 (2.9)</td>
<td>0.13 (2.7)</td>
<td>0.70 (1.3)</td>
<td>-0.08 (1.8)</td>
<td>84.0 (11.1)</td>
<td>77.74 (10.7)&lt;sup&gt;†&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

† = p < .10, * = p<.05, ** = p<.025, *** = p<.01, all two-tailed analyses

Note: The following comparisons were made, and the disjunctive response is italicized.  
Prisoner's Dilemma: 'Compete' vs. 'Cooperate'; Newcomb: 'Both Boxes' vs. 'Box B Only';  
Framing Problem: 'No Framing Effect' vs. 'Framing Effect'; Box Problem: 'No Preference'  
vs. 'Game 1'; Selection Task: 'Cor+P+All' vs. 'PQ'; Knights and Knaves: 'Knaves' vs. 'CBD'  
Charles in Scotland: 'Cor+Partial+June/Kate' vs. 'Charles in Scotland'; Married Problem  
and Green Problem: 'Yes' vs. 'CBD'
Table 17

Summary of Cognitive Ability and Thinking Disposition Differences For All Formal Decision-Making and Reasoning Problems

<table>
<thead>
<tr>
<th></th>
<th>Cognitive Ability</th>
<th>MFFT</th>
<th>AOT</th>
<th>Need for Cognition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Formal Decision-Making Problems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prisoner's Dilemma</td>
<td>Nonverbal 'Compete'</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Newcomb</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Framing Problem</td>
<td>--</td>
<td>--</td>
<td>'No Framing'</td>
<td>Composite, Total 'No Framing'</td>
</tr>
<tr>
<td>Box Problem</td>
<td>Verbal, Total 'No Preference'</td>
<td>Errors, RT-errors 'No Preference'</td>
<td>--</td>
<td>Composite, Total 'Game 2'</td>
</tr>
</tbody>
</table>

| **Formal Reasoning Problems** |                   |                       |                      |                        |
| Selection Task          | Total 'Cor+P+All'   | Errors, RT-errors 'Cor+P+All' | --                    | --                     |
| Knights & Knaves        | Nonverbal 'Knave'   | Errors 'Knave'        | 'Knave'              | Composite, Total       |
| Charles in Scotland     | Verbal, NonVerbal Total 'Cor+Par+J/K' | Errors, RT-errors 'Cor+Par+J/K' | --                   | --                     |
| Married Problem         | --                | --                    | --                   | --                     |
| Green Problem           | --                | --                    | --                   | Composite 'Yes'        |

Note 1: Cognitive Ability includes Verbal, Nonverbal, and Total Ability analyses; MFFT includes RT, Errors, and RT minus Errors; NeedCog reports effects with Composite and Total

Note 2: Responses recorded reflect higher scores, such as higher cognitive ability, higher scores on AOT composite, or fewer errors on MFFT.
### Table 18

**Cognitive Ability and Thinking Disposition Comparisons on The Prisoner's Dilemma Problem**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Compete (n=66) (Disjunctive)</th>
<th>Cooperate (n=57) (Non-Disjunctive)</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive Ability Measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Ability</td>
<td>-0.03 (1.75)</td>
<td>0.06 (1.93)</td>
<td>-0.27</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>0.34 (1.70)</td>
<td>-0.39 (1.71)</td>
<td>2.36**</td>
</tr>
<tr>
<td>Total Ability</td>
<td>0.31 (2.55)</td>
<td>-0.30 (3.03)</td>
<td>1.21</td>
</tr>
<tr>
<td><strong>Dispositional Measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOT Composite</td>
<td>0.02 (2.84)</td>
<td>0.13 (3.19)</td>
<td>-0.21</td>
</tr>
<tr>
<td>NeedCog Composite</td>
<td>-0.17 (1.81)</td>
<td>0.33 (1.81)</td>
<td>-1.54</td>
</tr>
<tr>
<td>NeedCog Scale Total</td>
<td>77.50 (10.8)</td>
<td>79.50 (10.3)</td>
<td>-1.05</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;RT&lt;/sub&gt;</td>
<td>0.00 (1.06)</td>
<td>0.01 (0.94)</td>
<td>-0.07</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;Errors&lt;/sub&gt;</td>
<td>0.07 (0.97)</td>
<td>-0.10 (1.01)</td>
<td>0.94</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;RT-Errors&lt;/sub&gt;</td>
<td>-0.07 (1.79)</td>
<td>0.11 (1.63)</td>
<td>-0.58</td>
</tr>
</tbody>
</table>

** = p < .025, all two-tailed, df=121

Note: AOT = Actively Open Minded Thinking, NeedCog = Need for Cognition, MFFT<sub>RT</sub> = Total Reaction Time Standard Score on MFFT, MFFT<sub>Errors</sub> Z-score = Total Number of Errors Standard Score on MFFT, MFFT<sub>RT-Errors</sub> Z-score = RT minus Errors Z-score.

Expected trends did not emerge on these dispositions, and in fact, the trend was in the reverse direction for the NeedCog thinking disposition, as participants who selected the Compete response displayed lower scores on the NeedCog-Composite and NeedCog-Total scales.

Some of the other dispositions that were used in the battery yielded some interesting patterns which may help to sort out the trends in Table 18. For example, those who chose to cooperate (M=4.35, SD=2.57) had a higher score on the Paranormal scale than those who chose to compete [M=3.50, SD=1.96; t(123)=-2.08, p<.05]. Those who chose to cooperate (M=2.97, SD=1.79) had a
marginally higher score on the Prayer scale than those who chose to compete \( M=2.39, SD=1.66; t(123)=-1.83, .05<p<.10 \). Also, those who chose to cooperate \( M=3.05, SD=1.84 \) had a marginally higher score on the Religiosity scale than those who chose to compete \( M=2.55, SD=1.53; t(121)=-1.67, .05<p<.10 \). These additional dispositions provide further insights into understanding performance on the individual difference measures for the Prisoner's Dilemma Problem. Those who chose to cooperate endorsed items related to Paranormal beliefs, Religious beliefs, and the Prayer Scale. Cooperators also displayed a slight trend on the Need for Cognition Composite Scale, but this trend did not reach significance. Those who chose to compete had significantly higher Nonverbal Cognitive Ability scores. These findings together provide an interesting profile that deserves further consideration, and will be discussed later.

**Newcomb's Problem.** Table 19 displays the simple comparisons on Newcomb's problem.

On Newcomb's Problem, there were no differences on any of the cognitive ability measures. This pattern of findings does not support a trend of higher cognitive ability for two box choosers. There were no differences found with any of the thinking dispositional measures. Those who chose the disjunctive two box response had a slightly higher score on the AOT composite than those who chose the non-disjunctive one box response, but this pattern did not approach significance. Overall, few results were obtained with Newcomb's Problem,

**Disease Framing Problem.** The cognitive ability and dispositional analyses are presented in Table 20 comparing those who displayed a framing effect with those who did not display a framing effect.

There were no significant relationships with any of the cognitive ability
Table 19

Cognitive Ability and Thinking Disposition Comparisons On Newcomb's Problem

<table>
<thead>
<tr>
<th>Variable</th>
<th>Both Boxes (n=46) (Disjunctive)</th>
<th>B Only (n=79) (Non-Disjunctive)</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Ability Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Ability</td>
<td>-0.25 (1.88)</td>
<td>0.14 (1.78)</td>
<td>-1.15</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>0.05 (1.73)</td>
<td>-0.03 (1.73)</td>
<td>0.24</td>
</tr>
<tr>
<td>Total Ability</td>
<td>-0.16 (2.57)</td>
<td>0.12 (2.88)</td>
<td>-0.53</td>
</tr>
<tr>
<td>Dispositional Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOT Composite</td>
<td>0.29 (2.59)</td>
<td>-0.12 (3.20)</td>
<td>0.74</td>
</tr>
<tr>
<td>NeedCog Composite</td>
<td>0.01 (1.95)</td>
<td>0.04 (1.78)</td>
<td>-0.09</td>
</tr>
<tr>
<td>NeedCog Scale Total</td>
<td>77.41 (11.85)</td>
<td>78.66 (10.09)</td>
<td>-0.62</td>
</tr>
<tr>
<td>MFFT_R T</td>
<td>0.07 (0.87)</td>
<td>-0.04 (1.07)</td>
<td>0.61</td>
</tr>
<tr>
<td>MFFT_Errors</td>
<td>0.08 (1.06)</td>
<td>-0.05 (0.97)</td>
<td>0.68</td>
</tr>
<tr>
<td>MFFT_R T-Errors</td>
<td>0.01 (1.65)</td>
<td>0.01 (1.79)</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

df=123

Note: AOT = Actively Open Minded Thinking, NeedCog = Need for Cognition, MFFT_R T Z-score = Total Reaction Time Standard Score on MFFT, MFFT_Errors Z-score = Total Number of Errors Standard Score on MFFT, MFFT_R T-Errors Z-score = RT minus Errors Z-score.

measures. However, those who displayed a framing effect had lower nonverbal and total ability scores, but these differences did not approach significance. Stanovich (1999; Stanovich & West, 1998c) reported a significant cognitive ability difference for the Disease Framing Problem on self-reported SAT scores. These results are not consistent with those reported by Stanovich (1999; Stanovich & West, 1998c). It is important to note that Stanovich reported this data from a larger sample (275 participants), which may account for the discrepancy between the results reported by Stanovich (1999; Stanovich & West,
Table 20

Cognitive Ability and Thinking Disposition Comparisons On Disease Framing Problem

<table>
<thead>
<tr>
<th>Variable</th>
<th>Framing Effect (n=26)</th>
<th>No Framing Effect (n=93)</th>
<th>F-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive Ability Measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Ability</td>
<td>-0.1 (1.78)</td>
<td>0.04 (1.88)</td>
<td>-0.33</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>-0.46 (1.46)</td>
<td>0.15 (1.76)</td>
<td>-1.62</td>
</tr>
<tr>
<td>Total Ability</td>
<td>-0.56 (2.08)</td>
<td>0.20 (2.97)</td>
<td>-1.22</td>
</tr>
<tr>
<td><strong>Dispositional Measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOT Composite</td>
<td>-0.94 (2.80)</td>
<td>0.34 (3.00)</td>
<td>-1.94*</td>
</tr>
<tr>
<td>NeedCog Composite</td>
<td>-0.67 (1.58)</td>
<td>0.32 (1.81)</td>
<td>-2.52**</td>
</tr>
<tr>
<td>NeedCog Scale Total</td>
<td>74.27 (10.01)</td>
<td>80.00 (10.31)</td>
<td>-2.52**</td>
</tr>
<tr>
<td>MFFTR T</td>
<td>0.09 (1.03)</td>
<td>-0.01 (1.02)</td>
<td>0.46</td>
</tr>
<tr>
<td>MFFTERrros</td>
<td>0.03 (1.03)</td>
<td>-0.00 (0.98)</td>
<td>0.16</td>
</tr>
<tr>
<td>MFFT R T-Errors</td>
<td>0.06 (1.72)</td>
<td>-0.01 (1.75)</td>
<td>0.18</td>
</tr>
</tbody>
</table>

* = p < .05, ** = p < .025, all two-tailed, df=117

Note: AOT = Actively Open Minded Thinking, NeedCog = Need for Cognition, MFFTR T-z-score = Total Reaction Time Standard Score on MFFT, MFFTERrros z-score = Total Number of Errors Standard Score on MFFT, MFFT R T-Errors z-score = RT minus Errors z-score.

1998c) and those reported in the present investigation.

Those who displayed no framing effect had a significantly higher score on the AOT disposition than those who displayed a framing effect. Those who displayed no framing effect also had a significantly higher score on the NeedCog Composite and on the NeedCog Raw Score Total than those participants who displayed a framing effect. The NeedCog effects replicate those found by Smith and Levin (1996). No relationships were obtained with the MFFT measures.
To summarize, cognitive ability differences were not obtained with the Disease Framing Problem, but some dispositional differences were found. The replication of Smith and Levin's (1996) findings with the Need for Cognition Scale increases confidence that this effect is robust. The dispositional differences include significant effects with the Need for Cognition Scale and Composite, and the AOT Composite. This combination of dispositions suggests that successful performance is not simply attributable to problem engagement, as captured by the Need for Cognition scale, but is also related to epistemic types of dispositions, such as actively open-minded thinking.

Overall, there are two interesting findings with the Disease Framing Problem: first, dispositional differences were found without cognitive ability differences, and second, dispositions related to problem engagement and actively open-minded thinking were associated with optimal performance on this problem.

**Box Problem.** Table 21 displays comparisons between Game 1 choosers, Game 2 choosers, and no preference choosers on the individual difference measures.

As shown in Table 21, no preference choosers had a significantly higher Total Cognitive Ability score and a significantly higher Verbal Ability score than both Game 1 and Game 2 choosers.

Some interesting patterns emerged on the Need for Cognition Composite and the Need for Cognition Total scores. That is, Game 2 choosers had a higher score on the NeedCog Composite than both no preference choosers and Game 1 choosers. Similarly, Game 2 choosers had a higher score on the NeedCog Total than both no preference choosers and Game 1 choosers. This was a somewhat unexpected result, as it was expected that no preference choosers would have higher scores in the Need for Cognition scale. This expectation was based on the assumption that those who realize the
### Table 21

#### Cognitive Ability and Thinking Disposition Comparisons On Box Problem

<table>
<thead>
<tr>
<th>Variable</th>
<th>Game 1 (n=84)</th>
<th>Game 2 (n=11)</th>
<th>No Preference (n=30)</th>
<th>F-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive Ability Measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Ability</td>
<td>-0.19 (1.77)</td>
<td>-0.64 (1.38)</td>
<td>0.77 (1.91)</td>
<td>3.98**ab</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>-0.05 (1.63)</td>
<td>-0.38 (1.94)</td>
<td>0.29 (1.89)</td>
<td>0.75</td>
</tr>
<tr>
<td>Total Ability</td>
<td>-0.22 (2.47)</td>
<td>-1.02 (2.72)</td>
<td>1.06 (3.30)</td>
<td>3.36*ab</td>
</tr>
<tr>
<td><strong>Thinking Disposition Measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOT Composite</td>
<td>-0.23 (2.83)</td>
<td>1.05 (3.35)</td>
<td>0.41 (3.25)</td>
<td>1.21</td>
</tr>
<tr>
<td>NeedCog Composite</td>
<td>-0.21 (1.78)</td>
<td>1.51 (1.87)</td>
<td>0.15 (1.77)</td>
<td>4.61**bc</td>
</tr>
<tr>
<td>NeedCog Scale Total</td>
<td>76.76 (10.74)</td>
<td>88.09 (8.34)</td>
<td>78.60 (9.81)</td>
<td>5.86***bc</td>
</tr>
<tr>
<td>MFFT RT</td>
<td>0.00 (1.03)</td>
<td>-0.49 (0.91)</td>
<td>0.18 (0.89)</td>
<td>1.81</td>
</tr>
<tr>
<td>MFFT Errors</td>
<td>-0.02 (0.995)</td>
<td>0.79 (1.09)</td>
<td>-0.24 (0.86)</td>
<td>4.54**bc</td>
</tr>
<tr>
<td>MFFT RT-Errors</td>
<td>0.02 (1.77)</td>
<td>-1.28 (1.66)</td>
<td>0.42 (1.47)</td>
<td>4.04**bc</td>
</tr>
</tbody>
</table>

* = p < .05, ** = p < .025, *** = p < .01, all two-tailed, df=117

Note 1: a = Game 1 vs. No Preference difference, b = Game 2 vs. No Preference difference, c = Game 1 vs. Game 2 difference
Note 2: AOT = Actively Open Minded Thinking, NeedCog = Need for Cognition, MFFT\text{RT-Z-score} = Total Reaction Time Standard Score on MFFT, MFFT\text{Errors-Z-score} = Total Number of Errors Standard Score on MFFT, MFFT\text{RT-Errors-Z-score} = RT minus Errors Z-score.

...equivalence of the options would be indifferent about which game they would choose, and these individuals would also be those who would have become more fully engaged in the problem, as captured by the NeedCog scale. However, Game 2 choosers are those who have a preference for a more complex event, and this may still be consistent with the NeedCog profile, as those high in NeedCog do tend to prefer more abstract and complex ideas. There were no differences approaching significance on the AOT Composite,
however, the AOT composite displays the same trend as the NeedCog measures, with Game 2 choosers having the highest mean score.

The MFFT measures displayed somewhat different patterns with Game 2 choosers. Game 2 choosers had significantly more errors on the MFFT (MFFT$_{Errors}$) than both Game 1 and No Preference choosers. Similarly, Game 2 choosers had significantly lower scores on the reaction time and errors discrepancy (MFFT$_{RT-Errors}$) than both Game 1 and no preference choosers. The negative score of Game 2 choosers on the MFFT$_{RT-Errors}$ measure indicates that these responders had a low total reaction time and many total errors, suggesting that these responders were impulsive and uninhibited in their response selections on the MFFT.

The results of the cognitive ability and dispositional differences on the Box Problem are somewhat complex. First, no preference choosers scored highest on the Cognitive Ability measures, in particular, Verbal Ability and Total Ability. The cognitive ability effects in favour of no preference choosers are consistent with a priori expectations. The next set of findings with Game 2 choosers were unexpected. That is, Game 2 choosers scored high on the NeedCog disposition, but displayed low total reaction time and many errors on the MFFT. Game 2 choosers also had significantly lower Verbal and Total Cognitive Ability scores than both Game 1 and no preference choosers. The profile of Game 2 choosers seems counter intuitive - how could one endorse items related to intellectual engagement on the Need for Cognition scale, yet display low scores on cognitive ability and MFFT performance?

One possibility is that Game 2 choosers are a group who are displaying a socially desirable response - which may be reflected in a high Need for Cognition score and selection of Game 2 in the Box Problem. That is, there may be some social demand characteristics to display intellectual interest (for
example, on the need for cognition scale), as measured in a self-report scale, and to select complex over simple events (namely, to choose Game 2). This alternative was explored by examining scores on the Impression Management scale, which was also a part of the Thinking Dispositions Questionnaire. There was no significant trend for Game 2 choosers on this scale, which suggests that social desirability does not account for the NeedCog finding.

One other possible explanation is a unique profile for Game 2 choosers that is not necessarily counter intuitive. It is possible that Game 2 choosers display an interest or need for highly engaging, complex, and abstract ideas and activities, which is captured by the Need for Cognition disposition. It would also make sense that these same individuals would prefer a complex event (Game 2) over a simple event. However, it is not necessary that these individuals have high cognitive ability or display reflective tendencies in order to respond this way on the Box Problem or on the Need for Cognition Scale.

**Formal Reasoning Problems**

**Selection Task.** The cognitive ability and dispositional analyses comparing disjunctive (Cor+P+All) responders with P,Q responders are presented in Table 22 (see page 63 for a description of how comparison groups were constructed).

Both cognitive ability measures and the thinking dispositions measures displayed effects in the expected directions. The Cor+P+All group had a significantly higher Total Cognitive Ability score than those who selected P,Q. The Verbal and Nonverbal ability scores was also marginally higher for the Cor+P+All group.

The Cor+P+All group had a significantly higher score on the AOT composite than those who selected P,Q. The Cor+P+All group had significantly fewer errors on the MFFT and a significantly higher score on
Table 22

Cognitive Ability and Thinking Disposition Comparisons On The Selection Task

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cor+P+All (n=58)</th>
<th>P,Q (n=50)</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Ability Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Ability</td>
<td>0.33 (1.89)</td>
<td>-0.33 (1.79)</td>
<td>1.87†</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>0.42 (1.65)</td>
<td>-0.16 (1.75)</td>
<td>1.79†</td>
</tr>
<tr>
<td>Total Ability</td>
<td>0.75 (2.69)</td>
<td>-0.46 (2.82)</td>
<td>2.29**</td>
</tr>
<tr>
<td>Dispositional Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOT Composite</td>
<td>0.74 (2.62)</td>
<td>-0.56 (3.24)</td>
<td>2.29**</td>
</tr>
<tr>
<td>NeedCog Composite</td>
<td>0.21 (1.92)</td>
<td>-0.15 (1.75)</td>
<td>1.01</td>
</tr>
<tr>
<td>NeedCog Scale Total</td>
<td>79.76 (11.04)</td>
<td>76.98 (10.48)</td>
<td>1.34</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;RT&lt;/sub&gt;</td>
<td>0.17 (1.01)</td>
<td>-0.07 (1.04)</td>
<td>1.24</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;Errors&lt;/sub&gt;</td>
<td>-0.33 (0.84)</td>
<td>0.23 (1.12)</td>
<td>-2.94***</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;RT-Erros&lt;/sub&gt;</td>
<td>0.50 (1.61)</td>
<td>-0.30 (1.87)</td>
<td>2.39***</td>
</tr>
</tbody>
</table>

† = p < .10,  ** = p < .025,  *** = p < .01, all two-tailed, df=106
Note: AOT = Actively Open Minded Thinking, NeedCog = Need for Cognition, MFFT<sub>RT</sub> Z-score = Total Reaction Time Standard Score on MFFT, MFFT<sub>Errors</sub> Z-score = Total Number of Errors Standard Score on MFFT, MFFT<sub>RT-Errors</sub> Z-score = RT minus Errors Z-score.

The MFFT<sub>RT-Errors</sub> measure than those who selected P,Q.

All of the results obtained on the Selection Task were in the expected direction. Both cognitive ability and thinking disposition measures (AOT, MFFT<sub>Errors</sub>, and MFFT<sub>RT-Errors</sub>) favoured the disjunctive response.

Selection Task: P Only Choosers versus P,Q Choosers. As P only choosers displayed a unique role in the domain generality analyses, P only choosers were examined separately in the same set of analyses with cognitive ability and thinking dispositions measures. The total cognitive ability effect was replicated in these analyses. The effect with AOT disappeared, but there
Disjunctive Thinking

was a trend in favour of P only choosers. The effects with MFFT_{Errors} were marginally in favour of P only choosers, and the MFFT_{RT-Errors} effect disappeared, but still displayed a trend in favour of P only choosers. Overall, the results of these analyses with P only choosers are similar to the results found with the P+Cor+All composite. These results further support the idea that choosing P only is an optimal choice.

**Knights and Knaves Problem.** The cognitive ability and thinking disposition comparisons for 'knave' and 'cannot be determined' responders are presented in Table 23.

On the cognitive ability measures, there were no significant differences on Total Cognitive Ability, but a trend was present in favour of knave responders. Knave responders had significantly higher Nonverbal Ability scores than those who selected the cannot be determined response.

The expected pattern emerged with the NeedCog disposition. Those who selected the knave response had a higher NeedCog Composite score than than those who selected the cannot be determined alternative. Also, those who selected the knave response had a higher NeedCog Raw Total Score than those who selected the cannot be determined alternative. The MFFT_{Errors} scores were consistent with both the Cognitive Ability and NeedCog patterns. That is, knave responders had significantly fewer total errors on the MFFT than the cannot be determined responders. No significant effects were found with MFFT_{RT} and MFFT_{RT-Errors}. There were no significant differences or notable patterns with the AOT Composite Scale. Therefore, both cognitive ability and some thinking disposition measures (NeedCog and MFFT_{Errors}) were implicated in performance on the Knights and Knaves Problem.

**Charles in Scotland Problem.** Cognitive ability and thinking dispositional measures were compared with broadly analytic versus heuristic
### Table 23

**Cognitive Ability and Thinking Disposition Comparisons on Knights and Knaves Problem**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Knave Response (n=52)</th>
<th>Cannot be Determined (n=61)</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive Ability Measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Ability</td>
<td>0.03 (1.80)</td>
<td>0.19 (1.91)</td>
<td>-0.47</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>0.51 (1.64)</td>
<td>-0.22 (1.68)</td>
<td>2.33**</td>
</tr>
<tr>
<td>Total Ability</td>
<td>0.57 (2.78)</td>
<td>-0.03 (2.68)</td>
<td>1.17</td>
</tr>
<tr>
<td><strong>Dispositional Measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOT Composite</td>
<td>0.13 (3.08)</td>
<td>0.23 (2.74)</td>
<td>-0.19</td>
</tr>
<tr>
<td>NeedCog Composite</td>
<td>0.56 (1.96)</td>
<td>0.29 (1.62)</td>
<td>2.53***</td>
</tr>
<tr>
<td>NeedCog Scale Total</td>
<td>81.25 (11.28)</td>
<td>76.69 (9.79)</td>
<td>2.30**</td>
</tr>
<tr>
<td>MFFTRT</td>
<td>0.01 (1.01)</td>
<td>0.02 (0.93)</td>
<td>-0.04</td>
</tr>
<tr>
<td>MFFTErrors</td>
<td>-0.31 (0.88)</td>
<td>0.19 (0.98)</td>
<td>-2.87***</td>
</tr>
<tr>
<td>MFFTRT-Errors</td>
<td>0.32 (1.64)</td>
<td>-0.18 (1.65)</td>
<td>1.61</td>
</tr>
</tbody>
</table>

** = p < .025, *** = p < .01, all two-tailed, df=111

Note: AOT = Actively Open Minded Thinking, NeedCog = Need for Cognition, MFFTRTZ-score = Total Reaction Time Standard Score on MFFT, MFFTErrorsZ-score = Total Number of Errors Standard Score on MFFT, MFFTRT-ErrorsZ-score = RT minus Errors Z-score.

responders, highly analytic versus undifferentiated other responders, and broadly analytic versus differentiated other responders. These results are displayed in Tables 24, 25, and 26. The broadly analytic versus heuristic responders comparison in Table 24 is discussed. The overall trends displayed in Tables 25 and 26 mirror the results in Table 24.

As displayed in Table 24, both Cognitive Ability and Thinking Dispositions differences were obtained for the Charles in Scotland Problem. Those who gave a broadly analytic response (Correct+Partial+June/Kate) had
Table 24

Cognitive Ability and Thinking Disposition Differences Between Broadly Analytic (Cor+Partial+June/Kate) and Heuristic (Charles) responders On The Charles in Scotland Problem

<table>
<thead>
<tr>
<th>Variable</th>
<th>Broadly Analytic (n=46)</th>
<th>Heuristic Responders (n=27)</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Ability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Ability</td>
<td>0.45 (1.56)</td>
<td>-0.51 (1.55)</td>
<td>2.52**</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>0.70 (1.63)</td>
<td>-1.02 (1.44)</td>
<td>4.54***</td>
</tr>
<tr>
<td>Total Ability</td>
<td>1.15 (2.54)</td>
<td>-1.53 (2.40)</td>
<td>4.43***</td>
</tr>
<tr>
<td>Thinking Dispositions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOT Composite</td>
<td>0.07 (2.77)</td>
<td>-0.77 (3.12)</td>
<td>1.19</td>
</tr>
<tr>
<td>NeedCog Composite</td>
<td>0.54 (1.86)</td>
<td>0.04 (1.48)</td>
<td>1.18</td>
</tr>
<tr>
<td>NeedCog Scale Total</td>
<td>81.2 (11.15)</td>
<td>77.78 (8.40)</td>
<td>1.39</td>
</tr>
<tr>
<td>MFFTRT</td>
<td>0.25 (0.98)</td>
<td>-0.19 (1.04)</td>
<td>1.83†</td>
</tr>
<tr>
<td>MFFTErrors</td>
<td>-0.20 (0.87)</td>
<td>0.51 (1.20)</td>
<td>-2.93***</td>
</tr>
<tr>
<td>MFFTRT-Errors</td>
<td>0.45 (1.57)</td>
<td>-0.71 (1.88)</td>
<td>2.76***</td>
</tr>
</tbody>
</table>

† = p < .10, * = p < .05, ** = p < .025, *** = p < .01, all two-tailed, df=71

Note: AOT = Actively Open Minded Thinking, NeedCog = Need for Cognition, MFFTRT Z-score = Total Reaction Time Standard Score on MFFT, MFFTErrors Z-score = Total Number of Errors Standard Score on MFFT, MFFTRT-Errors Z-score = RT minus Errors Z-score.

NOTE: Following response choices were excluded from the analysis: Other incorrect, Nothing, Blank.

a significantly higher Total Cognitive Ability score than those who responded with the heuristic response. These effects were also found for both Verbal and Nonverbal Ability.

No effects were found with the AOT Composite or with the Need for Cognition scales. However, significant effects were found with the MFFT. In particular, the broadly analytic responders had significantly fewer errors on the MFFT than heuristic responders. Then, the broadly analytic responders
Table 25

Cognitive Ability and Thinking Disposition Differences Between Highly Analytic (Cor+Partial) and Undifferentiated Other (June/Kate+Charles+Nothing+Blank+Other) Responders on the Charles in Scotland Problem

<table>
<thead>
<tr>
<th>Variable</th>
<th>Highly Analytic (n=31)</th>
<th>Undifferentiated Other (n=94)</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive Ability</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Ability</td>
<td>0.39 (1.53)</td>
<td>-0.13 (1.90)</td>
<td>1.38</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>0.63 (1.70)</td>
<td>-1.21 (1.69)</td>
<td>2.39**</td>
</tr>
<tr>
<td>Total Ability</td>
<td>1.02 (2.63)</td>
<td>-0.32 (2.74)</td>
<td>2.38**</td>
</tr>
<tr>
<td><strong>Thinking Dispositions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOT Composite</td>
<td>-0.24 (2.98)</td>
<td>0.12 (3.00)</td>
<td>-0.58</td>
</tr>
<tr>
<td>NeedCog Composite</td>
<td>0.47 (1.70)</td>
<td>-0.19 (1.87)</td>
<td>1.55</td>
</tr>
<tr>
<td>NeedCog Scale Total</td>
<td>80.26 (10.61)</td>
<td>77.52 (10.75)</td>
<td>1.23</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;RT&lt;/sub&gt;</td>
<td>0.38 (1.03)</td>
<td>-0.13 (0.96)</td>
<td>2.49**</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;Errors&lt;/sub&gt;</td>
<td>-0.25 (0.86)</td>
<td>0.08 (1.03)</td>
<td>-1.61</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;RT-Errors&lt;/sub&gt;</td>
<td>0.63 (1.58)</td>
<td>-0.21 (1.74)</td>
<td>2.37**</td>
</tr>
</tbody>
</table>

* = p < .05, ** = p < .025, *** = p < .01, all two-tailed, df=123

Note: AOT = Actively Open Minded Thinking, NeedCog = Need for Cognition, MFFT<sub>RT</sub> Z-score = Total Reaction Time Standard Score on MFFT, MFFT<sub>Errors</sub> Z-score = Total Number of Errors Standard Score on MFFT, MFFT<sub>RT-Errors</sub> Z-score = RT minus Errors Z-score.

had a significantly higher score on the MFFT<sub>RT-Errors</sub> measure than those who gave an heuristic response. A marginal effect in the same direction was also found on the MFFT<sub>RT</sub> measure.

Differences were also found on the Categorical Thinking Scale. Those who gave an heuristic response (M=7.33, SD=3.39) had a significantly higher score on this scale than those who gave a broadly analytic response [M=5.85, SD=2.12; t(71)=-2.31, p<.02]. Although differences on the AOT Composite were not found on this particular problem, the Categorical Thinking Scale is a
Table 26

<table>
<thead>
<tr>
<th>Variable</th>
<th>Broadly Analytic (n=46)</th>
<th>Differentiated Other (n=79)</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Ability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Ability</td>
<td>0.45 (1.56)</td>
<td>-0.26 (1.92)</td>
<td>2.11*</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>0.70 (1.63)</td>
<td>-0.41 (1.65)</td>
<td>3.64***</td>
</tr>
<tr>
<td>Total Ability</td>
<td>1.15 (2.54)</td>
<td>-0.65 (2.69)</td>
<td>3.66***</td>
</tr>
<tr>
<td>Thinking Dispositions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOT Composite</td>
<td>0.07 (2.77)</td>
<td>0.01 (3.12)</td>
<td>0.10</td>
</tr>
<tr>
<td>NeedCog Composite</td>
<td>0.54 (1.86)</td>
<td>-0.27 (1.77)</td>
<td>2.40**</td>
</tr>
<tr>
<td>NeedCog Scale Total</td>
<td>81.20 (11.15)</td>
<td>76.44 (10.16)</td>
<td>2.44**</td>
</tr>
<tr>
<td>MFFT_RT</td>
<td>0.25 (0.98)</td>
<td>-0.15 (0.99)</td>
<td>2.18*</td>
</tr>
<tr>
<td>MFFT_Errors</td>
<td>-0.20 (0.87)</td>
<td>0.12 (1.06)</td>
<td>-1.71</td>
</tr>
<tr>
<td>MFFT_RT-Errors</td>
<td>0.45 (1.57)</td>
<td>-0.26 (1.78)</td>
<td>2.25**</td>
</tr>
</tbody>
</table>

* = p < .05, ** = p < .025, *** = p < .01, all two-tailed, df=129

Note: AOT = Actively Open Minded Thinking, NeedCog = Need for Cognition, MFFT_RT Z-score = Total Reaction Time Standard Score on MFFT, MFFT_Errors Z-score = Total Number of Errors Standard Score on MFFT, MFFT_RT-Errors Z-score = RT minus Errors Z-score.

related disposition that pertains to epistemic regulation. This implicates the role of epistemic styles of thinking on performance in the Charles in Scotland Problem. Also, the cognitive ability and thinking dispositions differences mirror the findings on the Selection Task and the Knights and Knaves Problem.

**Married Problem.** The cognitive ability and dispositional analyses comparing Yes and Cannot be Determined responders are presented in Table 27.
Table 27

Cognitive Ability and Thinking Disposition Differences on The Married Problem

<table>
<thead>
<tr>
<th>Variable</th>
<th>Yes Response (n=16)</th>
<th>Cannot be Determined (n=107)</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Ability Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Ability</td>
<td>-0.68 (1.73)</td>
<td>0.11 (1.82)</td>
<td>-1.63</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>0.07 (1.71)</td>
<td>0.11 (1.72)</td>
<td>0.13</td>
</tr>
<tr>
<td>Total Ability</td>
<td>-0.61 (2.89)</td>
<td>0.14 (2.70)</td>
<td>-1.02</td>
</tr>
<tr>
<td>Dispositional Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOT Composite</td>
<td>-0.68 (3.46)</td>
<td>0.17 (2.89)</td>
<td>-1.07</td>
</tr>
<tr>
<td>NeedCog Composite</td>
<td>0.40 (1.99)</td>
<td>-0.04 (1.83)</td>
<td>0.88</td>
</tr>
<tr>
<td>NeedCog Scale Total</td>
<td>81.19 (10.13)</td>
<td>77.85 (10.88)</td>
<td>1.15</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;RT&lt;/sub&gt;</td>
<td>-0.28 (1.08)</td>
<td>0.04 (0.99)</td>
<td>-1.17</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;Errors&lt;/sub&gt;</td>
<td>0.05 (1.13)</td>
<td>0.01 (0.98)</td>
<td>0.12</td>
</tr>
<tr>
<td>MFFT&lt;sub&gt;RT-Errors&lt;/sub&gt;</td>
<td>-0.32 (1.96)</td>
<td>0.02 (1.71)</td>
<td>-0.74</td>
</tr>
</tbody>
</table>

all two-tailed, df=121

Note: AOT = Actively Open Minded Thinking, NeedCog = Need for Cognition, MFFT<sub>RT</sub> Z-score = Total Reaction Time Standard Score on MFFT, MFFT<sub>Errors</sub> Z-score = Total Number of Errors Standard Score on MFFT, MFFT<sub>RT-Errors</sub> Z-score = RT minus Errors Z-score.

As displayed in Table 27, there were no differences approaching significance on any of the cognitive ability or thinking dispositions measures for the Married Problem. In fact, there were trends in the alternative directions, as those who selected the yes response had lower average scores on the Verbal and Total Cognitive Ability measures than those who selected cannot be determined. The NeedCog Scale Total emerged in the expected pattern, with yes responders displaying a higher score on this scale than cannot be determined responders. The AOT and MFFT results showed the
reverse pattern, with yes responders scoring low on the AOT disposition and showing a lower reaction time on the MFFT metric.

**Green Levels Problem.** The individual difference analyses comparing Yes responders with Cannot be Determined responders is presented in Table 28.

On the cognitive ability measures, none of the relationships approached significant. Some interesting differences emerged with the NeedCog Composite, the NeedCog Scale Total, and MFFT_{Errors}. Those who responded disjunctively (‘yes’) on this problem had a higher score on the NeedCog Composite than non-disjunctive (‘cannot be determined’) responders. Then, disjunctive responders had a marginally higher score on the NeedCog Scale Total than non-disjunctive responders. Finally, disjunctive responders had marginally fewer errors on the MFFT than non-disjunctive responders. This pattern of findings on the Green Problem somewhat mirrors the pattern of results with the Knights and Knaves Problem, and further, provides some convergence on the domain generality found between the Green Levels Problem and the Knights and Knaves Problem.

Both the Married Problem and the Green Levels Problem originated from Levesque (1986, 1989), and these problems have the same embedded logic. The results for these two problems are complex, and there are a few consistencies. The disjunctive responders had a significantly higher Need for Cognition Composite score than non-disjunctive responders on the Green Levels Problem, and this trend was apparent with the Married Problem. The AOT Composite displayed a consistent direction in favour of non-disjunctive responders, but this was not significant for either of these problems. Subsequent work with these problems has helped to differentiate true ‘yes’ responders from chance ‘yes’ responders through the use of confidence.
Table 28
Cognitive Ability and Thinking Disposition Differences on The Green Levels Problem

<table>
<thead>
<tr>
<th>Variable</th>
<th>Yes Response (n=11)</th>
<th>Cannot be Determined (n=105)</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Ability Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Ability</td>
<td>0.76 (1.28)</td>
<td>0.03 (1.86)</td>
<td>1.28</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>0.27 (2.18)</td>
<td>0.08 (1.65)</td>
<td>0.35</td>
</tr>
<tr>
<td>Total Ability</td>
<td>1.03 (2.93)</td>
<td>0.13 (2.69)</td>
<td>1.05</td>
</tr>
<tr>
<td>Dispositional Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOT Composite</td>
<td>-0.53 (4.16)</td>
<td>0.1 (2.81)</td>
<td>-0.67</td>
</tr>
<tr>
<td>NeedCog Composite</td>
<td>1.13 (1.81)</td>
<td>-0.13 (1.79)</td>
<td>2.12*</td>
</tr>
<tr>
<td>NeedCog Scale Total</td>
<td>84.00 (11.05)</td>
<td>77.74 (10.67)</td>
<td>1.84†</td>
</tr>
<tr>
<td>MFFT_{RT}</td>
<td>0.21 (1.05)</td>
<td>-0.04 (0.98)</td>
<td>0.80</td>
</tr>
<tr>
<td>MFFT_{Errors}</td>
<td>-0.49 (0.78)</td>
<td>0.04 (1.00)</td>
<td>-1.71†</td>
</tr>
<tr>
<td>MFFT_{RT-Errors}</td>
<td>0.70 (1.25)</td>
<td>-0.08 (1.75)</td>
<td>1.44</td>
</tr>
</tbody>
</table>

† = p < .10, * = p < .05, all two-tailed, df=114
Note: AOT = Actively Open Minded Thinking, NeedCog = Need for Cognition, MFFT_{RT} Z-score = Total Reaction Time Standard Score on MFFT, MFFT_{Errors} Z-score = Total Number of Errors Standard Score on MFFT, MFFT_{RT-Errors} Z-score = RT minus Errors Z-score.

ratings. The most interesting finding with the Married Problem and the Green Levels Problem is their relationship to the Knights and Knaves Problem in the domain generality analyses. Conceptually, these two problems seem to capture an underlying disjunctive, problem solving process similar to what is necessary to solve the Knights and Knave Problem, and there is some support for this in the domain generality analyses.
Summary of Individual Difference Measures and Regression Analyses

In summary, some broad patterns can be extracted from the detailed analyses of trends in the individual difference measures on the formal tasks. These differences are summarized in Tables 17 and 18 (see pages 92-93).

The lack of differences with the Newcomb Problem, and the somewhat inconsistent findings with the Prisoner's Dilemma suggest that these problems do not fit well in the same domain or category as the other formal problems in the current study. The domain generality relationships found with the Married Problem and the Green Levels Problem suggest that these problems at least partly capture the notion of disjunctive thinking that was sought in this study. However, the very low frequency of disjunctive response selections on these problems, and the somewhat inconsistent differences on the individual difference analyses make these problems good candidates for further study.

The remaining formal problems are the Disease Framing Problem, the Box Problem, the Knights and Knaves Problem, the Selection Task, and the Charles in Scotland Problem. The individual difference measures have revealed some important differences and trends in the expected directions. That is, both cognitive ability and dispositional differences were obtained in favour of disjunctive response selections. Overall, these results suggest that computational power (cognitive ability), response regulation (as measured by the MFFT), and intellectual engagement (as measured by the Need for Cognition scale) are involved in explaining performance on some or all of these problems. One of the interesting questions is the separability of cognitive ability and thinking disposition measures. That is, do thinking dispositions predict performance after cognitive ability has been factored out? If unique variance is associated with both cognitive ability and thinking dispositions,
this would highlight the importance of both constructs in explaining reasoning performance. In order to address this question, a series of simultaneous multiple regressions was performed. Four regression analyses were performed with the following composite variables serving as criterion variables:

- Selection Task + Charles in Scotland + Knights and Knaves Problem (Sel+Char+Kni) performance (1);
- Box Problem performance (2);
- Knights and Knaves Problem performance (3); and
- Box Problem + Disease Framing Problem (Box+Framing) performance (4).

The composition and selection of these criterion variables is based on the a priori categorizations of formal decision-making and reasoning problems and on the summary of results presented in Table 17. Regression analyses with (1) and (3) include the formal reasoning problems, and analyses (2) and (4) include the formal decision-making problems. Also, based on the results presented in Table 17, cognitive ability and dispositional (specifically, Need for Cognition Composite and MFFT performance) differences were associated with performance on the Box Problem (2) and the Knights and Knaves Problem (3). Analyses (2) and (3) provided the purest test to examine the unique contributions of cognitive ability and thinking dispositions, as both cognitive ability and dispositions were implicated in performance on these two problems, as displayed in Table 17. The composite variables (1) and (4) were formed in order to examine whether there was any generality within the different sets of formal problems with respect to the individual difference measures. The composition of composites (1) and (4) are supported by the trends and differences displayed in Table 17. For the regression analyses (1)
and (4), cognitive ability was combined as a predictor with either MFFT_{Errors} or the Need for Cognition thinking disposition. The results of these regression analyses are displayed in Table 29.

The results of the regression analyses and the commonality analyses displayed in Table 29 provide evidence that both cognitive ability and thinking dispositions explain performance on the formal tasks, and in fact, both contribute unique variance in explaining performance. For example, MFFT_{Errors} explains unique variance on the formal reasoning problems (Sel+Char+Kni). Further, MFFT_{Errors} explains 7.4% of the variance, which is as much variance as accounted for by cognitive ability. This finding is exciting in light of questions that have been raised about the relationship between cognitive ability and MFFT performance (Salkind & Wright, 1977), as it highlights the stylistic aspect of this measure. However, MFFT_{Errors} did not significantly contribute additional variance to performance on the formal decision-making Box Problem.

The Need for Cognition disposition also yielded some interesting effects in the regression analyses. On the Knights and Knaves Problem, Total Cognitive Ability did not predict performance, but the Need for Cognition disposition significantly explained performance on this problem. On the Box and Framing Problems, both Total Cognitive Ability and Need for Cognition were significant predictors. What is noteworthy in both of these analyses is the separability of thinking dispositions and cognitive ability, and that the dispositional variable of Need for Cognition accounts for nearly as much (Box+Framing) or more variance (Knights and Knaves Problem) than is attributable to cognitive ability.

Informal Decision-Making Task

The Consumer Task. Correlational analyses were performed in order to
Table 29

Simultaneous Regression Analyses With Formal Decision-Making and Reasoning Problems

<table>
<thead>
<tr>
<th>Criterion Variable</th>
<th>Variable</th>
<th>beta weight (standardized)</th>
<th>F-value</th>
<th>Unique Variance Explained</th>
<th>Partial r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sel+Char+Kni</td>
<td>Cognitive Ability</td>
<td>0.284</td>
<td>3.361***</td>
<td>0.074</td>
<td>0.292</td>
</tr>
<tr>
<td></td>
<td>MFFT Errors</td>
<td>-0.285</td>
<td>3.366***</td>
<td>0.074</td>
<td>-0.293</td>
</tr>
<tr>
<td>Overall Regression:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F= 16.131***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple R = 0.459</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple R-squared = 0.211</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Box Problem</td>
<td>Cognitive Ability</td>
<td>0.192</td>
<td>2.01*</td>
<td>0.035</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>MFFT Errors</td>
<td>-0.058</td>
<td>0.605</td>
<td>0.004</td>
<td>-0.057</td>
</tr>
<tr>
<td>Overall Regression:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F= 2.652 †</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple R = 0.214</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple R-squared = 0.046</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knights and Knaves</td>
<td>Cognitive Ability</td>
<td>0.061</td>
<td>0.646</td>
<td>0.003</td>
<td>0.062</td>
</tr>
<tr>
<td>Problem</td>
<td>Need for Cognition</td>
<td>0.22</td>
<td>2.321**</td>
<td>0.046</td>
<td>0.216</td>
</tr>
<tr>
<td>Overall Regression:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F= 3.398*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple R = 0.241</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple R-squared = 0.058</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Box Problem + Framing Problem</td>
<td>Cognitive Ability</td>
<td>0.186</td>
<td>2.092*</td>
<td>0.033</td>
<td>0.186</td>
</tr>
<tr>
<td></td>
<td>Need for Cognition</td>
<td>0.174</td>
<td>1.951*</td>
<td>0.028</td>
<td>0.174</td>
</tr>
<tr>
<td>Overall Regression:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F= 5.268***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple R = 0.282</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple R-squared = 0.079</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† = p < .10,  * = p < .05, ** = p < .025, *** = p < .01
explore relationships between performance on the Consumer Task and the individual difference measures. These analyses are displayed in Tables 30, 31, 32, and 33. Overall, very few significant correlations were obtained in these analyses. While some speculations can be made about the patterns that have emerged in the data, I hesitate to interpret these patterns because of the possibility of capitalizing on chance.

No relationships between cognitive ability and the number of unique features generated for each of the items nor for the total features generated were obtained (see Table 30). Some relationships were found with the thinking dispositions scales (see Table 31). It was found that those who generated more unique features for the bread maker had a higher score on the Need for Cognition Composite scale. Two relationships with $MFFT_{\text{Errors}}$ were obtained. That is, those who generated more unique features for the car and those who generated more total features tended to display fewer errors on the MFFT. In sum, the few effects that were obtained were in the anticipated direction, but they did not emerge systematically across all of the items.

It is interesting that no relationships were obtained between the Consumer Task and cognitive ability, but some small relationships in the anticipated direction were obtained with the thinking dispositions. In Table 30, many of the correlations are nearly zero, suggesting no relationship between features generated and cognitive ability. In Table 31, results with the MFFT measures display positive relationships between features generated and the $MFFT_{\text{RT}}$ and $MFFT_{\text{RT-Errors}}$ measures, and negative relationships with $MFFT_{\text{Errors}}$. Given the lack of systematicity and significant effects, these trends cannot be interpreted with any confidence.

Correlational analyses were also performed between matched features generated from Consumer Reports magazine and cognitive ability (Table 32),
Table 30

Correlations between Unique Features Generated and Cognitive Ability on Consumer Task

<table>
<thead>
<tr>
<th></th>
<th>Verbal Ability</th>
<th>Nonverbal Ability</th>
<th>Total Cognitive Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breadmaker</td>
<td>0.17</td>
<td>0.10</td>
<td>0.16</td>
</tr>
<tr>
<td>T-Shirt</td>
<td>0.03</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Car</td>
<td>-0.02</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>Total Features</td>
<td>0.07</td>
<td>0.05</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Note 1: Only the relevant correlations are displayed. Note 2: All ability measures refer to cognitive ability, and features from the Consumer Task are unique features generated.

Table 31

Correlations between Unique Features Generated and Thinking Dispositions on Consumer Task

<table>
<thead>
<tr>
<th></th>
<th>AOT</th>
<th>NeedCog</th>
<th>MFFTRT</th>
<th>MFFTErrors</th>
<th>MFFTRT-Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bread maker</td>
<td>0.09</td>
<td>0.20*</td>
<td>0.07</td>
<td>-0.09</td>
<td>0.08</td>
</tr>
<tr>
<td>T-Shirt</td>
<td>0.05</td>
<td>-0.04</td>
<td>0.03</td>
<td>-0.15</td>
<td>0.10</td>
</tr>
<tr>
<td>Car</td>
<td>0.01</td>
<td>0.00</td>
<td>0.15</td>
<td>-0.20*</td>
<td>0.20</td>
</tr>
<tr>
<td>Total Features</td>
<td>0.06</td>
<td>0.07</td>
<td>0.11</td>
<td>-0.18*</td>
<td>0.16</td>
</tr>
</tbody>
</table>

* p<.05  Note 1: Only the relevant correlations are displayed.
Note 2: AOT=Actively Open Minded Thinking, NeedCog=Need for Cognition, MFFTRT = standardized total reaction time to first response on Matching Familiar Figures Test, MFFTErrors = standardized total number of errors on Matching Familiar Figures Test, MFFTRT-Errors = standardized discrepancy score between total reaction time to first item and total number of errors, and features from Consumer Task are unique features generated.
and with thinking dispositions (Table 33). A somewhat different pattern emerged in these analyses. There was a significant relationship between the Consumer Reports features generated for the bread maker and Verbal Cognitive Ability, and between Consumer Reports features generated for the t-shirt and Verbal Cognitive Ability. These relationships were also obtained with Total Cognitive Ability for both the bread maker and the t-shirt. The positive direction of these relationships suggests that cognitive ability may be implicated in the Consumer Reports features that were generated. The Consumer Reports features are a selected subset of features that are intended to discriminate diagnostic from non-diagnostic aspects of products. The positive trend with cognitive ability displayed in Table 31 suggests that the Consumer Reports features are discriminative in a way that was not evident when the total unique features generated were correlated with the cognitive ability measures (see Table 30).

Some interesting patterns also emerged between the Consumer Reports features and the thinking dispositions (Table 33). The Consumer Reports features generated for the car were significantly and positively related to reaction time on the MFFT and the MFFT<sub>RT-Errors</sub> score. Finally, the total features generated from Consumer Reports was negatively related to the total number of errors on the MFFT. That is, those who generated more features that matched those listed by Consumer Reports had fewer errors on the MFFT. The trends displayed with the thinking dispositions in Table 33 are very similar to the patterns displayed in Table 31, when only unique features were considered.

Together, the Consumer Reports scoring provided a somewhat different picture with the individual difference measures, and the interpretation of the trends is only speculative. The generation of Consumer Reports features
Table 32

**Correlations between Consumer Reports Features Generated and Cognitive Ability on Consumer Task**

<table>
<thead>
<tr>
<th>Features</th>
<th>Verbal Ability</th>
<th>Nonverbal Ability</th>
<th>Total Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breadmaker</td>
<td>0.24*</td>
<td>0.07</td>
<td>0.20*</td>
</tr>
<tr>
<td>T-Shirt</td>
<td>0.23*</td>
<td>0.09</td>
<td>0.20*</td>
</tr>
<tr>
<td>Car</td>
<td>0.08</td>
<td>0.17</td>
<td>0.15</td>
</tr>
<tr>
<td>Total Features</td>
<td>0.08</td>
<td>0.06</td>
<td>0.09</td>
</tr>
</tbody>
</table>

*p<.05

All ability measures refer to cognitive ability, and features from Consumer Task are matched features from Consumer Reports magazine.

Table 33

**Correlations between Consumer Reports Features Generated and Thinking Dispositions on the Consumer Task**

<table>
<thead>
<tr>
<th>Features</th>
<th>AOT</th>
<th>NeedCog</th>
<th>MFFT&lt;sub&gt;RT&lt;/sub&gt;</th>
<th>MFFT&lt;sub&gt;Errors&lt;/sub&gt;</th>
<th>MFFT&lt;sub&gt;RT-Errors&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breadmaker</td>
<td>0.11</td>
<td>0.14</td>
<td>0.14</td>
<td>-0.15</td>
<td>0.17</td>
</tr>
<tr>
<td>T-Shirt</td>
<td>-0.04</td>
<td>0.04</td>
<td>0.11</td>
<td>-0.01</td>
<td>0.07</td>
</tr>
<tr>
<td>Car</td>
<td>-0.00</td>
<td>0.04</td>
<td>0.18*</td>
<td>-0.15</td>
<td>0.19*</td>
</tr>
<tr>
<td>Total Features</td>
<td>0.10</td>
<td>0.09</td>
<td>0.09</td>
<td>-0.19*</td>
<td>0.16</td>
</tr>
</tbody>
</table>

*p<.05

Note: AOT=Actively Open Minded Thinking, NeedCog=Need for Cognition, MFFT<sub>RT</sub>=standardized total reaction time to first response on Matching Familiar Figures Test, MFFT<sub>Errors</sub>=standardized total number of errors on Matching Familiar Figures Test, and MFFT<sub>RT-Errors</sub>=standardized discrepancy score between total reaction time to first item and total number of errors.
generated displayed trends with Cognitive Ability and Thinking Dispositions, in particular, on MFFT performance. Alternatively, when the total number of unique features generated were considered, only dispositions were implicated, in particular, MFFT performance.

The cognitive ability relationships obtained in the Consumer Reports generated features, but not with the total features generated is interesting. The Consumer Reports features are a list of those diagnostic features necessary to make a quicker and more efficient decision. The total unique features generated is a completely generative measure which captures the initial phases of disjunctive consideration, but does not embrace anything evaluative or diagnostic in the same manner as the Consumer Reports generated features. While this analysis lends credence, it provides no explanatory power as to why the Consumer Reports features should necessarily be positively correlated with cognitive ability.

The positive relationships between thinking dispositions obtained between both unique features generated and the Consumer Reports features generated suggests that something dispositional, perhaps separable from cognitive ability, is involved in the task of disjunctive generation. Simply engaging with the task and deriving a list of features requires some thoughtful consideration, and consequently, more features generated must be an index of at least longer consideration time. Further thoughtful consideration is also more likely to result in identifying at least some diagnostic features that would be listed by a source like Consumer Reports. The dispositional effects obtained were with the MFFT measures. The MFFT measures involve tendencies toward response regulation and reflectivity, and these tendencies seem to be good descriptions of performance on this task. With the few effects obtained, this interpretation is a conservative estimation of
the role of thinking dispositions in the Consumer Task.

One important consideration, which was also examined in this study, was the role of prior experience with purchasing the different products on Consumer Task performance.

**The Consumer Task and Prior Experience.** Prior experience was examined as an exploratory variable, as presumably, someone who has purchased a bread maker may be more likely to generate a different set of features than someone who has never purchased a bread maker. Participants were asked to report whether they had ever purchased a bread maker, a t-shirt, or a car. Eleven participants (8.8%) reported that they had previously purchased a bread maker. Then, 121 (96.8%) participants reported that they had previously purchased a t-shirt. Finally, 51 (40.8%) participants reported that they had previously purchased a car. Participants' experiences with each of these products were therefore, relatively different.

Some stepwise hierarchical regression analyses were performed in order to examine whether prior experience was a predictor of feature generation, and whether cognitive ability and thinking dispositions were predictors after prior experience had been partialled out. Prior experience was first entered as a predictor for each item, and then the different individual measures were also entered as predictors, in particular, Total Cognitive Ability, the AOT Composite, the NeedCog Composite, and the MFFT measures. Prior experience was only a significant predictor for the bread maker. No significant, consistent, or interpretable patterns with the individual difference measures emerged from these analyses.

**Informal Reasoning Task**

**Argument Generation Task.** The same statistical approach was used for this task as for the Consumer Task. Correlational analyses were
performed in order to explore relationships between performance on the Argument Generation Task, cognitive ability, and thinking dispositions. A series of analyses was also performed with each issue (tuition, organs, and gasoline) in relation to cognitive ability and thinking dispositions measures, but no systematic or explainable relationships emerged for an individual issue. Thus, for parsimony and clarity, the analyses reported here were done with composites of total arguments, my-side arguments, other-side arguments, and belief bias collapsed across all three issues. This approach is supported by the significant correlations between the total arguments generated for each issue in the domain generality analyses (see Table 12).

In Table 34, correlations between total number of arguments generated, total number of my-side arguments generated, total number of other-side arguments generated, and belief bias in relation to the cognitive ability measures are presented. Table 35 displays correlations between the same Argument Generation Task measures and thinking disposition measures. In Table 34, two relationships were obtained with the Verbal Ability measure. The total number of arguments generated and the total number of other-side arguments generated were positively correlated with Verbal Ability. No relationships were obtained between argument generation and the thinking dispositions measures (Table 35). What is interesting in these analyses is that Verbal Ability correlated with total arguments and other-side arguments, but NOT with my-side arguments. This finding is consistent with the previously reported findings on domain generality within the Argument Generation Task. That is, in the domain generality analyses, it was found that some domain generality was found between different issues, and between my-side and other-side arguments for different issues, BUT no relationship was found between my-side and other-side arguments within the same issue (see Table
### Table 34
**Correlations Between Argument Generation Task and Cognitive Ability**

<table>
<thead>
<tr>
<th></th>
<th>Verbal Ability</th>
<th>Nonverbal Ability</th>
<th>Total Cognitive Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Arguments</td>
<td>0.18*</td>
<td>-0.00</td>
<td>0.12</td>
</tr>
<tr>
<td>My-side Arguments</td>
<td>0.10</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>Other-side Arguments</td>
<td>0.18*</td>
<td>0.03</td>
<td>0.15</td>
</tr>
<tr>
<td>Belief Bias</td>
<td>-0.06</td>
<td>-0.03</td>
<td>-0.06</td>
</tr>
</tbody>
</table>

*p<.05* Note: Relevant correlations are bolded. All ability measures refer to cognitive ability, and Total Arguments = Total number of arguments generated for all 3 issues, My-side Arguments = Total number of my-side arguments generated for all 3 issues, Other-side Arguments = Total # of other side arguments generated for all 3 issues, and Belief Bias = total my-side minus total other side arguments.

### Table 35
**Correlations Between Argument Generation Task and Thinking Dispositions**

<table>
<thead>
<tr>
<th></th>
<th>AOT</th>
<th>NeedCog</th>
<th>MFFT&lt;sub&gt;RT&lt;/sub&gt;</th>
<th>MFFT&lt;sub&gt;Errors&lt;/sub&gt;</th>
<th>MFFT&lt;sub&gt;RT-Errors&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Arguments</td>
<td>-0.08</td>
<td>-0.00</td>
<td>0.08</td>
<td>-0.09</td>
<td>0.10</td>
</tr>
<tr>
<td>My-side Arguments</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.05</td>
<td>-0.10</td>
<td>0.08</td>
</tr>
<tr>
<td>Other-side Arguments</td>
<td>-0.13</td>
<td>-0.05</td>
<td>0.10</td>
<td>-0.05</td>
<td>0.09</td>
</tr>
<tr>
<td>Belief Bias</td>
<td>0.08</td>
<td>0.09</td>
<td>-0.04</td>
<td>-0.05</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Note 1: Only the relevant correlations are displayed
Note 2: AOT=Actively Open Minded Thinking, NeedCog=Need for Cognition, MFFT<sub>RT</sub>=standardized total reaction time to first response on Matching Familiar Figures, MFFT<sub>Errors</sub>=standardized total number of errors on Matching Familiar Figures, MFFT<sub>RT-Errors</sub>=standardized discrepancy score between total reaction time to first item and total number of errors.
13). The effect with cognitive ability here suggests the same separability between my-side and other-side arguments as displayed by the domain generality analyses. However, this interpretation is speculative, as the few relationships obtained may also be due to chance.

As was done with the Consumer Task, the role of prior information or experience was also examined for the Argument Generation Task by considering the impact of prior beliefs on argument generation for each issue.

**Argument Generation Task and Prior Belief.** The impact of prior belief has been demonstrated in some important studies of critical thinking (for example, Sá, West, & Stanovich, 1999; Stanovich & West, 1997). Prior belief can be efficacious in some task scenarios (Sá, West, & Stanovich, 1999), but it has also been associated with non-optimal reasoning (Stanovich, 1999).

A series of stepwise hierarchical regression analyses were performed in order to determine whether Prior Opinion Strength would be significant predictor of the number of arguments generated. Then, cognitive ability and dispositional measures were entered in order to determine whether they would predict the number of arguments generated after Prior Opinion Strength had been entered as a predictor variable. The outcome measures were the number of my-side and other-side arguments generated for each issue. Analyses were done separately for each issue because prior beliefs on each of the three issues varied considerably, particularly between the tuition issue and the other two issues. On the tuition issue, 65.6% (n=82) of participants endorsed having a Strong Prior Belief, 18.4% (n=23) endorsed having a Moderate Prior Belief, and 16% (n=20) endorsed having a Weak Prior Belief. On the organs issue, 35.2% (n=44) of participants endorsed having a Strong Prior Belief, 44.0% (n=55) endorsed having a Moderate Prior Belief, and 20.8% (n=26) endorsed having a Weak Prior Belief. On the gasoline issue,
36.8% (n=46) of participants endorsed having a Strong Prior Belief, 34.4% (n=43) endorsed having a Moderate Prior Belief, and 28.8% (n=36) endorsed having a Weak Prior Belief.

Overall, Prior Opinion Strength was found to be a significant predictor of the number of my-side arguments for the tuition issue, and a significant predictor of the number of other-side arguments for the gasoline issue. Cognitive ability was not a significant predictor in any of the analyses. In one instance, the AOT disposition was a significant predictor of the number of other-side arguments generated for the gasoline issue. Given the few effects and lack of systematicity in the effects obtained, these effects are likely due to chance. However, these results implicate the important role of prior beliefs on argument generation.

In addition to prior belief, some analyses were also done in order to examine the role of knowledge in argument generation.

**Argument Generation Task and Knowledge.** Recall that participants were asked some knowledge questions about tuition fees in Michigan, about the salaries of university graduates with Ph.D.'s (both pertaining to the tuition issue), and about gasoline prices in other countries (gasoline issue). Table 36 displays some important relationships that were obtained between Prior Opinion Strength, Knowledge, and Belief Bias.

The pattern of results presented in Table 36 display the Belief Bias Index as the criterion variable on two different dimensions: Level of Knowledge (Misconception or Correct) and Prior Opinion Strength (Strong, Medium, or Weak). Specifically, two knowledge questions pertaining to tuition costs at universities were asked, and one question pertaining to gasoline costs was asked. Some interesting trends emerge when knowledge was introduced as an additional variable. In particular, the results in Table 36 suggest that it is
Table 36

Belief Bias on The Argument Generation Task as a Function of Prior Opinion Strength and Knowledge

<table>
<thead>
<tr>
<th>Prior Opinion Strength</th>
<th>Strong</th>
<th>Medium</th>
<th>Weak</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tuition Issue: Knowledge of Tuition Fees in Michigan</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Misconception</td>
<td>0.80 (n=25)</td>
<td>2.00 (n=6)</td>
<td><strong>1.28</strong> (n=7)</td>
</tr>
<tr>
<td>Correct</td>
<td><strong>1.54</strong> (n=57)</td>
<td>0.88 (n=16)</td>
<td>0.31 (n=13)</td>
</tr>
<tr>
<td><strong>Tuition Issue: Knowledge of salaries of Ph.D. Graduates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Misconception</td>
<td>0.68 (n=22)</td>
<td>0.37 (n=27)</td>
<td>0.15 (n=20)</td>
</tr>
<tr>
<td>Correct</td>
<td><strong>1.30</strong> (n=23)</td>
<td>0.75 (n=16)</td>
<td>-0.31 (n=16)</td>
</tr>
<tr>
<td><strong>Gasoline Issue: Knowledge of Gasoline Prices in Other Countries</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Misconception</td>
<td><strong>1.44</strong> (n=18)</td>
<td>0.25 (n=4)</td>
<td><strong>1.13</strong> (n=8)</td>
</tr>
<tr>
<td>Correct</td>
<td><strong>1.30</strong> (n=63)</td>
<td>1.35 (n=17)</td>
<td>0.33 (n=12)</td>
</tr>
</tbody>
</table>

conceivable that someone can have a strong belief bias (as displayed by more my-side arguments and self-reported prior belief) and be well calibrated because their strong belief is based on some knowledge. This trend was evident on all three analyses (this is indicated by the bolded numbers in Table 36). There was also evidence of poorly calibrated belief bias, as evidenced by a lack of knowledge, more my-side arguments, and weak self-reported prior beliefs (numbers also bolded in Table 36). The role of knowledge displayed in Table 36 may explain the lack of effects between the thinking dispositions and the manner in which well-calibrated reasoning was measured in this study. While the number of my-side versus other-side arguments may be a partial index of disjunctive thinking, the role of knowledge, as displayed in Table 36, captures another aspect of disjunctive thinking which is clearly important,
but was only tangentially explored in this study. Knowledge may be a mediating variable that partially determines when belief bias is appropriate or misplaced. The results in Table 36 suggest that domain specific knowledge about an issue may be as important as considering both sides of the issue.

In summary, initial explorations between the Argument Generation Task and the individual difference measures yielded few of the anticipated patterns. Some effects were obtained with cognitive ability measures, and no relationships were obtained with any of the thinking dispositions. When prior beliefs on the issues were taken into account, it was found that there was large variation in participants' prior opinions for each of the three issues. This lack of systematicity of prior beliefs further substantiated my hesitation to interpret the few trends and lack of trends that were obtained with the individual difference measures. Some explorations with the knowledge questions that were used in this study implicated the important role of domain specific knowledge. In particular, different levels of knowledge may discriminate between poorly calibrated and well calibrated belief bias.

4. Other Notable Findings

Thinking Dispositions

Finally, a correlation matrix displaying relationships between the cognitive ability measures and the target thinking dispositions appears in Table 37. As can be seen in Table 37, MFFT performance clearly has a positive association with cognitive ability, confirming the source of the debate on whether MFFT performance is a measure of style or cognitive ability (Salkind & Wright, 1977). However, MFFT performance was not related to the Need for Cognition or Actively Open-Minded Thinking dispositions in this study, suggesting a separability between these dispositions. This is consistent with the treatment of the MFFT as a disposition related to response regulation, and
Table 37

Correlations Between Cognitive Ability and Thinking Dispositions Measures

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Verbal Ability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Nonverbal Ability</td>
<td>0.25**</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Total Ability</td>
<td>0.80***</td>
<td>0.78***</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. AOT Composite</td>
<td>0.18</td>
<td>0.09</td>
<td>0.17</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. NeedCog Composite</td>
<td>0.17</td>
<td>0.17</td>
<td>0.22**</td>
<td>0.36***</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. MFFT&lt;sub&gt;RT&lt;/sub&gt;</td>
<td>0.23**</td>
<td>0.01</td>
<td>0.15</td>
<td>-0.14</td>
<td>-0.05</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. MFFT&lt;sub&gt;Errors&lt;/sub&gt;</td>
<td>-0.23**</td>
<td>-0.25**</td>
<td>-0.30***</td>
<td>0.02</td>
<td>-0.03</td>
<td>-0.51***</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>8. MFFT&lt;sub&gt;RT-Errors&lt;/sub&gt;</td>
<td>0.27**</td>
<td>0.15</td>
<td>0.26**</td>
<td>-0.10</td>
<td>-0.01</td>
<td>0.87***</td>
<td>-0.87***</td>
<td>---</td>
</tr>
</tbody>
</table>

* = p<.05, ** = p<.025, *** = p<.01

Note: AOT=Actively Open-Minded Thinking, NeedCog=Need for Cognition Composite Scale, MFFT<sub>RT</sub>=MFFT Total Reaction Time to First Response, MFFT<sub>Errors</sub>=Total Errors, MFFT<sub>RT-Errors</sub>=Difference between Total Reaction Time to First Response and Total Number Errors on MFFT.

the Need for Cognition and AOT dispositions as tendencies toward engagement, persistence in thinking, and epistemic regulation. The scores on the Need for Cognition and the Actively Open-Minded Thinking dispositions were significantly correlated, suggesting that these two dispositions are highly related. The Need for Cognition disposition was also significantly correlated with the Total Cognitive Ability measure, but the Actively Open-Minded Thinking disposition was not correlated with any of the ability
measures. The Need for Cognition Composite, however, was not related to MFFT performance. MFFT performance and the Need for Cognition Composite displayed the most systematic effects on the formal problems. The lack of correlations between these two dispositions provides evidence that they are statistically separable and that they should be conceptualized as independent styles or tendencies.

**Concluding Discussion**

The results reported in this investigation provide support for a number of interesting conclusions about the domain of disjunctive thinking.

First, Shafir (1994) was clearly right about the disjunction effect. This study confirmed that participants have difficulty with solving tasks that involve disjunctive thinking. This was shown on a variety of problems and tasks from the formal decision-making and reasoning literatures. In addition, the larger proportion of my-side than other-side arguments generated in the Argument Generation Task was also evidence for a lack of disjunctive consideration.

The trends and differences observed in performance on the formal tasks suggest some domain generality of a disjunctive thinking skill, particularly in the formal reasoning tasks. However, the lack of associations between many of the formal reasoning and decision-making tasks provided stronger evidence for the specificity of performance. The results of the informal tasks displayed patterns that provided evidence for a different way of conceptualizing disjunctive thinking. There was evidence for a generational fluency in the informal tasks. That is, those who generated more features on the Consumer Task had a tendency to generate more arguments on the Argument Generation Task. These generative tendencies were also observed within each
of the informal tasks.

The pattern of findings obtained in the individual difference examinations provided useful trends for understanding who does well on disjunctive thinking problems. Regression analyses provided evidence for the separability and unique contributions of cognitive ability and thinking dispositions on the formal reasoning and decision-making tasks. Few associations were found between performance on the informal tasks and the individual difference measures. Some evidence suggested that the role of prior experience in the Consumer Task, and the roles of prior belief and knowledge in the Argument Generation Task may impact reasoning about specific issues differently. The Generic Dual Process Framework of reasoning from Stanovich (1999; Stanovich & West, 2000) was used to conceptualize the patterns of results found in this study.

Is Disjunctive Thinking A Domain General Skill?

Patterns of Associations on the Formal Reasoning Tasks

Shafir's (1994) theoretically-based categorization of disjunctive problems received some empirical support in this study. The results support some domain generality of disjunctive thinking in a subset of the formal reasoning and decision-making problems used in this study. Figure 6 displays a diagrammatic conceptualization of the general patterns of associations between the set of formal tasks used in this study. It is important to note that this visual display is not based on any statistical or multidimensional scaling techniques.

The pattern of associations displayed in Figure 6 provides some indications of a disjunctive thinking skill that, most specifically, connects the set of formal reasoning problems used in this study. In addition to Shafir (1994), there is other theoretical support that ties the formal reasoning
Figure 6. Diagrammatic Conceptualization of the Domain Generality Relations Between the Formal Tasks
problems together. Levesque (1986, 1989) described the Married Problem and the Green Levels Problem as logical puzzles. The Knights and Knaves Problem has been described as an instance of deductive reasoning, and was originally selected by Rips (1989) from a book of logic puzzles from Smullyan (1978). The Selection Task has often been described as a formal logic task that involves hypothesis testing (Baron, 1994). The Charles in Scotland Problem has been called a formal logic task that involves propositional reasoning (Johnson-Laird, Byrne, & Schaeken, 1992). The pattern of results in this study and the rich empirical history of these problems provide some support for a domain general skill of disjunctive thinking that is related to deductive logic and problem solving.

A less coherent relationship is evident among the formal decision-making problems, and between the formal reasoning and decision-making tasks. There is some support for a domain general problem solving component in the Disease Framing Problem and the Box Problem that is tied to realizing the equivalence of the choices in the problems. The Prisoner's Dilemma and Newcomb's Problem were, at best, related marginally to performance on one other formal problem. Overall, the patterns of performance on the formal decision-making tasks and between the formal reasoning and decision-making tasks more strongly support a domain specificity view of performance on the formal tasks that were used in this study.

What do we make of Newcomb's Problem and the Prisoner's Dilemma?

Newcomb. The version of Newcomb's problem used in the current study was analogous to the version discussed by Stanovich (1999). Originally from Nozick (1969), this version was adapted by Shafir and Tversky (1992; Shafir, 1994) in which the predictor's choice had already been made, and which removes the seemingly supernatural element that was present in Nozick's
Disjunctive Thinking

(1969) version. Stanovich (1999) used the Generic Dual Process Framework and the application of the understanding/acceptance principle to illustrate how the two box selection is an optimal solution (Shafir, 1994; Shafir & Tversky, 1992). Stanovich (1999) reported that those participants who selected the two box solution had higher scores on the Need for Cognition disposition. No significant differences in self-reported SAT scores were found by Stanovich (1999), but a higher SAT trend was apparent for the two box choosers.

Cognitive ability differences were not found between those who selected both boxes and those who selected the one box choice in this study, which is somewhat consistent with Stanovich (1999). Need for cognition differences were not found in this study, but were reported by Stanovich (1999). It is important to examine why the two box selection was not related to disjunctive response selections on the other formal tasks, and why the individual difference metrics did not correlate in favour of the two box response selection.

In Stanovich’s (1999) application of the understanding/acceptance principle, he reported that when participants were presented with a normative argument in favour of choosing both boxes or when participants were presented with a normative and non-normative argument, a significant number of non-normative choosers opted for the normative choice. When an argument was presented in favour of the non-normative Box B Only choice, a significant number of normative choosers switched in favour of the non-normative argument. It is noteworthy that the non-normative argument was so compelling that a significant number of participants’ switched their responses to the non-normative option — is this a reflection of an alternative task construal?

The philosophical literature also provides a possible explanation for selecting the less optimal Box B Only response. The dominance principle
would take the stance that one should take both boxes instead of only one box because the potential gain in both boxes is higher than what can be obtained from only one of the boxes (Nozick, 1993). However, Nozick (1993) also introduced the idea of causally based expected utilities. If any causally expected utility is invoked between the predictor and what is obtained in the boxes, that is, that there is "some causal-probabilistic relation indicating direct causal influence" (Nozick, 1993, p. 43), one could argue that the one box solution is the better choice. If a causally based expected utility is invoked, an individual would be more likely to select the box with the highest utility (in this case, Box B Only) based on reliance on the "Predictor". But as pointed out by Shafir (1994), reliance on the "Predictor" reflects a form of quasi-magical thinking, that some external controller has reliably predicted your behaviour. However, the argument posed by Nozick (1993) suggests an alternative task construal: that if a causally based interpretation is invoked (whether or not this construal is optimal), then the one box solution is the optimal choice.

These considerations highlight the complexities of solving Newcomb's Problem, and support the idea that this problem cannot be viewed as an instance that only requires disjunctive consideration. It is possible that one could have traveled all the branches of the decision tree in this problem, and yet still have arrived at a causally based interpretation that yields the one-box choice.

An interesting point made by Nozick (1993) is that the amount of the reward in Box A is an important cue:

By varying the amount of money in the first box, we can make people extremely uncomfortable with their otherwise favored arguments for choice in Newcomb's initial problem. People who initially choose both boxes are unwilling to follow the dominance argument when the amount in the first box is lowered to $1; people who initially chose only the second box are unwilling to follow the expected utility argument (with conditional probabilities that do not mark influence) when the
amount in the first box is raised to $900000. This suggests that no one has complete confidence in the argument he or she follows for Newcomb’s initial example. (p. 45)

One possible hypothesis does emerge from Nozick’s (1993) discussion which is relevant to Stanovich’s (1999) Generic Dual Process Framework for reasoning. That is, if the amount in Box A is increased substantially to greatly exceed the value in Box B, then the causally driven heuristic preference for the one box solution should be overridden by the dominance principle to select both boxes. That is, substantially increasing the amount in Box A should result in the cuing of both heuristic and analytic systems, resulting in the selection of the more optimal, both boxes solution in both instances. If such findings were obtained, they would then be analogous to Stanovich’s (1999) findings with the deontic version of the selection task.

The evidence suggests that performance on Newcomb’s problem is less related to performance on the other formal disjunctive problems in this study. The philosophical literature (Nozick, 1993) and the empirical findings reported by Stanovich (1999) support the possibility that there are alternative task construals for Newcomb’s problem. In particular, the finding reported by Stanovich (1999), that the non-normative argument compelled normative choosers to select the one box option suggests that the difficulty in this problem does not lie in reaching the end of each branch in this decision tree. Instead, the results in this investigation and these theoretical discussions support the idea that the difficulty in this problem may be that none of branches provide an undeniably optimal solution.

Prisoner’s Dilemma. Some questions have also been raised about the optimal solution in the Prisoner’s dilemma problem. In particular, the application of the dominance principle would favour the ‘compete’ option as
the optimal solution (Lewis, 1979), and the evidentially expected utility argument would favour the ‘cooperate’ option as the optimal response (Nozick, 1993).

According to the dominance principle, Lewis (1979) has argued that the Prisoner’s dilemma is analogous to Newcomb’s problem. Consider the following matrix:

<table>
<thead>
<tr>
<th></th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate</td>
<td>Cell A</td>
</tr>
<tr>
<td>You: 20</td>
<td>You: 5</td>
</tr>
<tr>
<td>Other: 20</td>
<td>Other: 25</td>
</tr>
<tr>
<td>Compete</td>
<td>Cell C</td>
</tr>
<tr>
<td>You: 25</td>
<td>You: 10</td>
</tr>
<tr>
<td>Other: 5</td>
<td>Other: 10</td>
</tr>
</tbody>
</table>

According to the dominance principle, you should prefer Cell C to Cell A and Cell D to Cell B, while the other should prefer Cell B to Cell A and Cell D to Cell C. In an evidentially expected utility argument, some would argue that people may bring implicit assumptions into their reasoning, in particular, that whatever they reason themselves, the other will behave in the same rational manner (Nozick, 1993). According to Nozick’s (1993) evidentially expected utility argument, the following assumptions are made:

1. each agent has common knowledge
2. each is a rational agent
3. each prefers a to d

Based on the third assumption, an evidentially expected utility argument would suggest that to cooperate is the normative choice. Shafir (1994) argues that if one relies on a purely disjunctive approach, the compete response selection is the optimal choice. However, an evidential approach is an alternative task construal that reduces the certainty of the compete option as
the optimal choice in this problem.

Another possible interpretation is based on the trends found with the individual difference measures in this investigation, which also supports an alternative task construal position. Cooperators displayed a slight trend on the Need for Cognition Composite Scale, but those who chose to compete had significantly higher Nonverbal Cognitive Ability scores. The Need for Cognition trend was not in the expected direction, which suggests that arriving at the choice to compete may not be a purely disjunctive process - via reflectivity, engagement, and consideration of all of the alternatives. This leaves open the possibility that one may have considered both alternatives thoughtfully in this problem, and still have arrived at the non-disjunctive option.

Also on the thinking disposition measures, it was found that those who chose to cooperate endorsed items related to Paranormal beliefs, Religious beliefs, and endorsed the item on the Prayer Scale. The trends on the Religiosity and Prayer Scale dispositions implicate an adoption of moral principles, for example, avoiding the choice of gains at another’s expense. The trend on the Paranormal Scale is a reflection of what Shafir and Tversky (1992) have called ‘magical’ thinking, which refers to “the erroneous belief that one can influence an outcome by some symbolic or other indirect act even though the act has no causal link to the outcome” (p. 463). Interestingly, Shafir and Tversky (1992) attribute the selection to cooperate on the Prisoner’s Dilemma to ‘quasi-magical’ thinking which refers to cases in which “people act as if they erroneously believe that their action influences the outcome, even though they do not really hold that belief” (p. 463). Shafir and Tversky (1992) make a distinction between ‘magical’ and ‘quasi-magical’ thinking, but report that both are puzzling because ‘magical’ thinking commits one to absurd beliefs,
and 'quasi-magical' thinking undermines the link between belief and action. These two types of thinking may not be so different, and those who endorsed items on the Paranormal scale display beliefs, whether real or quasi, that they may have some control over the outcome. Consistent with the analysis of Newcomb's problem, the philosophical literature and the data in this study support alternative task construals in the Prisoner's dilemma problem.

Patterns of Associations on the Informal Reasoning Tasks

Evidence for 'Generational Fluency.' A consistent set of positive correlations was found with the Consumer Task and the Argument Generation Task. To summarize, the following relationships were obtained in this investigation:

- the number of unique features generated for each item in the Consumer Task were positively correlated;
- the total number of arguments generated for each issue in the Argument Generation Task were positively correlated; and
- the total number of unique features generated in the Consumer Task were positively related to the total number of arguments generated in the Argument Generation Task.

These results provide evidence for a domain general skill, but this skill is different from the type of disjunctive consideration that was described previously in the formal reasoning and decision-making problems. This skill may be conceptualized as a generational fluency, which is likely related to disjunctive thinking, but does not necessarily involve fully traveling all the branches of the decision tree as in the formal problems. Something akin to generational fluency or disjunctive generation would explain the bulk of these relationships. That is, those who have a tendency to generate more features for the bread maker are the same individuals who would generate more
arguments for the organ issue. It is important to emphasize the pure
generation aspect that is intended in this description of generational fluency,
as it is different than how others have conceptualized generation in other
contexts of reasoning. For example, Torrens, Thompson, and Cramer (1999)
used an alternatives generation task in which participants had to generate
diagrams based on syllogisms -- there is no model being elaborated in the way
the term generational fluency is described here. The informal tasks used in
this study had few constraints, demanding a skill related primarily to
generation.

Where There Was No Generational Fluency. Generational fluency was
a consistent trend in both the Consumer Task and the Argument Generation
Task. However, one trend was found in which there was no trend in
generation, signaling a disjunction effect in the Argument Generation Task.
In particular, no relationship was found between the numbers of my-side and
other-side arguments that were generated within each issue. For example,
participants who provided more my-side arguments for the tuition issue
provided fewer other-side arguments for the tuition issue. This pattern was
found for all three issues. These lack of correlations can be interpreted as
evidence for a disjunction effect, or a failure to search for arguments on both
sides (Baron, 1995). That is, there is evidence for a poor balance between
equating or considering the two sides of a given issue. These effects replicate
those reported by Baron (1995), as Baron also reported no significant
correlations between the number of my-side and other-side arguments
generated when he asked participants to reason about the abortion issue.

Baron (1994) also reported some interesting evidence related to
preferences for one-sided thinking. That is, participants rated one-sided
thinking better than two-sided thinking. How can one-sided thinking be
advantageous if a lack of considering alternatives, or the disjunction effect, is associated with less optimum reasoning in this study? Baron states:

One possibility is that people confuse good thinking with expertise. Experts do not need to think, and consideration of the other side suggests that expertise is lacking. It is also possible that certain institutions, such as organized religions, promote the idea that seeing two sides is “confusing”. (p. 233).

Baron (1994) argues that two-sided thinking is associated with an active search for reasons why an idea might be wrong and with “actively open-mindedness”. Is two-sided thinking really better than one-sided thinking?

Some further explorations with prior belief and domain knowledge measures provided further insight into the disjunction effect in the Argument Generation Task. Domain knowledge was found to discriminate between two different patterns of belief bias, as measured by the discrepancy between the number of my-side and other-side arguments generated. The two patterns were: domain knowledge, strong prior beliefs, and belief bias tendencies versus no domain knowledge, weak prior beliefs, and belief bias tendencies. These are two very different profiles, but the interesting finding demonstrated in this study is that knowledge may be an important moderator of belief bias tendencies -- unfounded belief bias versus belief bias reflecting well calibrated thinking. While belief biases are typically associated with poor reasoning (Baron, 1995; Klaczynski, Gordon, & Fauth, 1997; Klaczynski & Narasimham, 1998), some research has suggested that it can be advantageous to rely on prior beliefs in some cases (Sá, West, & Stanovich, 1999). In this case, someone who can back up their strong beliefs with domain knowledge should have more my-side arguments available than other-side arguments. The results of the current study suggest that knowledge is domain specific, and having some knowledge may help for a particular issue but does not likely transfer to other
issues. The domain specificity of knowledge in reasoning has also been suggested in other research (Perkins, Farady, & Bushey, 1991; Lewandowsky & Kirsner, 2000).

**Summary: Domain Generality and Disjunctive Thinking in Informal and Formal Tasks**

Investigation of the domain generality hypothesis provided some evidence for a generational fluency in the informal tasks and evidence for some domain generality between the formal reasoning problems. Few relationships were obtained between the formal decision-making and reasoning problems. No relationships were found between response tendencies on the informal tasks and the formal problems. This dissociation between the informal and formal tasks suggests that disjunctive thinking in these two broad sets of tasks, as measured in this study, are unrelated. For example, realizing that individual C can only be a ‘knave’ in the knights and knaves problem is unrelated to the number of features generated in the Consumer Task. It is insufficient in the knights and knaves problem to simply generate the possible alternatives, namely, to realize that individual C could be a knight or a knave. In the Consumer Task, there are substantially more alternatives available to generate than in the knights and knaves problem, but it is only this process of generation that seems to be analogous in these two tasks. The positive correlations within the Consumer Task, for example, provide evidence for some type of generative skill, but this skill is unrelated to the type of disjunctive consideration that is captured by the formal tasks.

The distinction between generation and evaluation has been described by Baron (1994) in a search-inference framework. Baron (1994) distinguished between possibilities, goals, and evidence. Possibilities are possible answers to the original question. Goals are the criteria by which possibilities are
Disjunctive Thinking

Evidence consists of any belief, or potential beliefs, that helps one determine the extent to which a possibility can achieve some goal. The important distinction from these results is between possibilities versus goals and evidence. The informal tasks entailed the generation of possibilities, while the formal tasks (in particular, the reasoning tasks) set the expectation of deriving a single solution, which would have involved progression towards this goal via use of evidence and information in the problems.

**Who Does Well on Disjunctive Thinking Problems?**

**Formal Reasoning and Decision-Making Problems**

Systematic patterns in cognitive ability and thinking dispositions were found to predict performance on the formal reasoning problems. In fact, both the Need for Cognition Composite and MFFT performance were found to contribute unique variance after cognitive ability was statistically controlled. Table 38 displays each formal problem and displays the cognitive ability and thinking dispositions findings obtained for each problem. Responses on each of the problems were collapsed into the categories of ‘heuristic response’ and ‘analytic failure/success’ following Stanovich (1999). The main trends displayed in Table 38 can be summarized as follows:

1. Consistent with the domain generality analyses, many of the differences and trends obtained with the individual difference measures occurred with the formal reasoning problems. Few relationships were obtained between the individual difference measures and the formal decision-making problems.
2. Both cognitive ability and thinking disposition differences were obtained on the formal reasoning and decision-making problems. Simultaneous hierarchical regressions were performed, and it was found that both cognitive capacities and thinking dispositions predicted
<table>
<thead>
<tr>
<th>Decision Making Problems</th>
<th>Heuristic Response</th>
<th>Analytic Failure/Success</th>
<th>Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newcomb</td>
<td>Box B Only</td>
<td>Both Boxes</td>
<td>AOT trend</td>
</tr>
<tr>
<td>Prisoner’s Dilemma</td>
<td>Cooperate</td>
<td>Compete</td>
<td>Nonverbal Ability</td>
</tr>
<tr>
<td>Disease Framing Problem</td>
<td>Framing Effect</td>
<td>No Framing Effect</td>
<td>AOT NeedCog</td>
</tr>
<tr>
<td>Box Problem</td>
<td>Game 1</td>
<td>Game 2</td>
<td>Cognitive Ability</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reasoning Problems</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Knights and Knaves Problem</td>
<td>Cannot be determined</td>
<td>Knave</td>
<td>Nonverbal Ability NeedCog MFFT Errors</td>
</tr>
<tr>
<td>Married Problem</td>
<td>Cannot be determined</td>
<td>Yes</td>
<td>NeedCog trend</td>
</tr>
<tr>
<td>Green Levels Problem</td>
<td>Cannot be determined</td>
<td>Yes</td>
<td>NeedCog MFFT Errors trend</td>
</tr>
<tr>
<td>Selection Task</td>
<td>P,Q</td>
<td>Cor+P+All</td>
<td>Cognitive Ability AOT MFFT Errors MFFT RT-Errors</td>
</tr>
<tr>
<td>Charles in Scotland</td>
<td>Charles in Scotland</td>
<td>Cor+Par+June/Kate</td>
<td>Cognitive Ability MFFT Errors MFFT RT-Errors</td>
</tr>
</tbody>
</table>
unique variance for performance on some of these tasks. The main thinking dispositions that were associated with performance on the formal disjunctive tasks were the Matching Familiar Figures Test and the Need for Cognition measures.

The lack of associations observed between the MFFT and the Need for Cognition scale were somewhat unexpected. It was expected that actively open-minded thinking (AOT), need for cognition, and reflectivity-impulsivity (as measured by the MFFT) would be overlapping constructs, and that they would all explain performance on the disjunctive tasks used in this study. This expectation was consistent with other conceptualizations of the links between these dispositions (Baron, 1994; Stanovich, 1999). For example, Baron (1994) suggested that reflectives were more likely actively open-minded thinkers than impulsives. In this study, \( MFFT_{\text{Errors}} \) was a more sensitive and discriminative measure than \( MFFT_{\text{RT}} \) or the \( MFFT_{\text{RT-Errors}} \) measures. It was expected that the MFFT reaction time measure would measure the types of tendencies that are captured by the AOT and the Need for Cognition scale. The AOT and Need for Cognition scales include tendencies such as intellectual engagement, intellectual curiosity, persistence, thinking about alternatives, and a general interest in thinking. There was an overall dissociation between AOT and \( MFFT_{\text{Errors}} \) and the Need for Cognition scale and \( MFFT_{\text{Errors}} \) across the the formal tasks, which suggests that \( MFFT_{\text{Errors}} \) captures something unique from the Need for Cognition and AOT scales. Someone who has made many errors on the MFFT is not only displaying impulsivity, but is also displaying other tendencies, such as failing to double check their answer in the face of other very similar options. Frequent errors on the MFFT also seem to characterize a tendency to give a response without
being absolutely certain that the response is correct. These tendencies captured by \( \text{MFFT}_{\text{Errors}} \) may be termed as a type of response regulation. In a review of executive function measures and developmental psychopathology, Pennington and Ozonoff (1996) reported that \( \text{MFFT}_{\text{Errors}} \) more often discriminated between ADHD and control samples. As ADHD is viewed as a disorder of disinhibition (Barkley, 1998), this conceptualization of \( \text{MFFT}_{\text{Errors}} \) as a measure of response regulation is consistent with other research. Therefore, the results of the present investigation suggest that \( \text{MFFT}_{\text{Errors}} \) capture something distinct from the types of tendencies captured by the Actively Open-Minded Thinking Scale and the Need for Cognition Scale.

Finally, the unique variance accounted for by \( \text{MFFT}_{\text{Errors}} \) after cognitive ability has been partialled out provides strong evidence for the stylistic aspect of the \( \text{MFFT}_{\text{Errors}} \) measure, which has been a source of much debate in the developmental literature (Egeland & Weinberg, 1976; Block, Block, & Harrington, 1974). Overall, these findings provide a broader conceptualization of stylistic elements that predict reasoning performance. Stanovich (1999) has primarily studied thinking dispositions which are related to epistemic regulation, such as on tasks of argument evaluation (Stanovich & West, 1997) and belief bias (Sá, West, & Stanovich, 1999). The style of reflectivity-impulsivity, as measured by the MFFT, continues in the the tradition of identifying malleable predictors of reasoning performance, and adds another facet to understanding the nature of these stylistic tendencies and how they are related to reasoning performance. Therefore, this study provides evidence to suggest that both epistemic regulation (as measured by the Need for Cognition scale) and response regulation (as measured by the MFFT) are both important predictors of disjunctive thinking performance. Both of these categories of dispositions provide different conceptualizations of how thinking
performance can be altered -- that is, through intellectual engagement and intellectual curiosity AND by spending more time on a problem and monitoring performance.

The Special Case of P Only Choosers in the Selection Task. The original work on the Selection Task has regarded the selection of the cards P and Not-Q as the ‘correct’ solution (Evans, Newstead, & Byrne, 1993). Margolis (1987) made a strong case for treating P only choosers as an optimal selection under an alternative task construal. Margolis suggests that in an open reading of the task, categories are chosen instead of cards. In the rule: “If a card has a vowel on its letter side, then it has an even number on its number side”, an open reading would invoke checking the category of vowels or the category of odd numbers in order to determine whether the rule is broken. According to Margolis (1987), it is unnecessary to check both categories, as this would be redundant. Some evidence has been found to support this construal (see Evans, Newstead, & Byrne, 1993 for a review).

Stanovich (1999; Stanovich & West, 1998a) integrated Margolis’ interpretation into his classification of response patterns on the selection task into an heuristic-analytic framework. Heuristic processing is the relatively fast and automatic response selection, while analytic processing is slower and more controlled. On non-deontic versions of the selection task, Stanovich identified correct responders as those who have attained analytic success. He termed P only choosers, P, Q, Not-Q choosers, and All choosers as those who have experienced analytic failure. Finally, P,Q choosers, or the ‘matching’ responders, were called heuristic responders. Stanovich and West (1998a; also reported in Stanovich, 1999) report cognitive ability differences which discriminate analytic success, analytic failure, and heuristic responders. As reported in Stanovich and West (1998a), heuristic responders (P, Q) had lower
self-reported SAT score than those who displayed analytic failure (P only; P, Not-Q; and All). Then, those who experienced analytic failure had a lower self-reported SAT score than those who experienced analytic success (P, Not-Q). This evidence suggests that P only and All choosers demonstrate higher level reasoning than heuristic P, Q choosers. The P, Q response has been referred to as the "matching bias" in the cognitive psychology literature (Evans, Newstead, & Byrne, 1993), and this term succinctly captures the heuristic tendency to select those responses which match the exemplars stated in the rule (Stanovich, 1999; Evans, 1995). The results of this study, previous literature by Margolis (1987), and Stanovich (1999; Stanovich & West, 1998a) all provide a compelling argument for including the selection of P only choosers as also an optimal choice.

One issue that is somewhat puzzling in this data is that sometimes P only choosers appeared more disjunctive than those who selected the correct P, Not-Q response. For example, some of the domain generality relationships between the selection task and other formal tasks were obtained only with P only choosers; these relationships were obtained with the knights and knaves problem, the Disease framing problem, and the Prisoner’s dilemma. There are two possible directions in which to interpret these results: either to promote P only choosers from ‘analytic failure’ to ‘analytic success’ status, or to place less emphasis on the domain general relations obtained with P only choosers (as two of the three relations obtained were with formal decision-making problems, which yielded few relations with the other problems). Both interpretations are appropriate and would be supported by the data reported in this study. Conceptualizations of P only choosers as an analytic-type task construal (Margolis, 1987; Stanovich & West, 1998; Stanovich, 1999) is more strongly supported by data presented in this investigation.
Explaining Performance on the Informal Tasks

The individual difference measures provided few patterns to help understand performance on the informal tasks. On the Argument Generation Task, these results did not replicate those reported by Perkins, Farady, and Bushey (1991). Perkins et al. (1991) reported that high IQ was positively correlated with the number of my-side arguments generated, but they did not find such a relationship between IQ and number of other-side arguments generated. In the present study, no consistent relationships across the three issues were found with cognitive ability, and no differences were obtained with the thinking dispositions measures. Importantly, Perkins et al. (1991) reported this effect with only one issue, therefore it is difficult to interpret the lack of replication in the current study. Similarly on the Consumer Task, Verplanken (1993) reported a significant relationship between performance on an informal decision-making task and the Need for Cognition disposition. This effect was also not replicated in this investigation with the Consumer Task, which was modeled on Verplanken's task. While it was disappointing not to replicate Perkins et al. (1991) and Verplanken (1993), these results do not deny the importance of cognitive ability and thinking dispositions on performance in the informal tasks. Rather, the evidence reported here implicated the important role of prior experience on the Consumer Task, and prior beliefs and knowledge on the Argument Generation Task. It is feasible that there are systematic relationships between prior beliefs and disjunctive performance on the Argument Generation Task, however, the large differences in prior beliefs between the issues studied here prevent any systematic investigation of the impact of knowledge and prior beliefs on belief bias tendencies. The lack of relations between the belief bias indices on the three issues and the impact of knowledge displayed in this study supports the
domain specificity of performance on the Argument Generation Task. An extreme stance of domain specificity in knowledge (for example, Alexander, Kulikowich, & Schulze, 1994; Ceci, 1993, 1996), however, may not be appropriate. There may be cases in which belief bias on the Argument Generation Task is domain general, as has been reported in previous research (Sá, West, & Stanovich, 1999). For example, knowledge may be domain specific, but an individual having equivalent knowledge and prior beliefs on a set of issues should be better calibrated on those issues than someone who has little knowledge and no prior beliefs on the same set of issues.

Why do People Fail to Reason Disjunctively?

Implications for Models of Human Reasoning

The pattern of findings in this investigation are best conceptualized under Stanovich's (1999) Generic Dual Process Framework for reasoning performance. The association between optimal or disjunctive response selections and trends or differences displayed with the individual differences investigated support this model of reasoning (see Table 38). First, the patterns displayed in Table 38 strongly support a differentiation between heuristic and analytic response tendencies on the formal disjunctive reasoning problems, and some support on the formal decision-making problems. Second, the cognitive ability and thinking dispositions differences provide further conceptualizations of System 2 processes. The findings reported in this investigation also converge with other theories and perspectives on reasoning, including heuristic-analytic models of reasoning (Evans, 1984, 1989), and Damasio's somatic marker hypothesis (1994).

What new knowledge have we gained about heuristic and analytic processes from the set of formal disjunctive decision-making and reasoning tasks that were examined in this study? The categorical differentiation
between heuristic and analytic solutions on the formal problems displayed in Table 38 is supported by previous conceptualizations of these problems in the literature, and with the current findings of this study. For example, the P,Q response in the Selection Task has been described as an heuristic response in the literature (Stanovich, 1999; Evans, 1995). The ‘matching bias’ descriptive (Evans, Newstead, & Byrne, 1993) captures well the easiest route to a solution: pick the cards indicated in the question because they must be relevant, else they would not be mentioned (Grice, 1975). Use of a relevance heuristic fits well with the automatic default strategy described as System 1 processing (Stanovich, 1999). The Charles in Scotland problem can be interpreted in a similar manner. The Charles in Scotland response was associated with lower cognitive ability scores than responding ‘nothing’ or leaving the question blank. Therefore, the Charles in Scotland response may reflect a tendency or desire to give an answer, and the shortest route to derive an answer may be to extract the most commonly stated proposition in the question, that ‘Charles is in Scotland.’ The ‘nothing’ response or leaving the question blank may reflect a slightly higher order response, such as an inhibition of relying on such a relevance heuristic. On the other formal reasoning problems (knights and knaves problem, married problem, green levels problem), the cognitive ability and thinking dispositions differences were also in favour of the disjunctive solutions. This pattern of findings suggests that the ‘cannot be determined’ response option also reflects an heuristic response  

One interesting consideration is that the heuristic and analytic responses may be highly contingent on the response options. That is, one might argue that to respond with ‘nothing’ in the Charles in Scotland problem is analogous to a ‘cannot be determined’ response option in the other problems. However, the fact that the correct answer is available to select in the knights and knaves, married, and green levels problems constrains the heuristic and analytic responses to the options presented. A good example to illustrate this possibility is the Box Problem from Shafir (1994): the addition of the ‘no preference’ option in this study created a new analytic response option.
The responses selections on the formal reasoning problems can be grouped as analytic or heuristic with relative ease under Stanovich's (1999) Generic Dual Process Framework, but this discrimination is more difficult with the formal decision-making problems. Damasio's (1994) discussion of the role of emotion maps well onto System 1 processes in Stanovich's (1999) framework, and this link is useful for discriminating heuristic and analytic responses selections in at least some of the decision-making problems. In particular, the Disease Framing Problem and the Prisoner's Dilemma.

Damasio (1994) provides a fitting description of the role of emotion in rational decision making. He states:

I see some failures of rationality as not just due to a primary calculation weakness, but also due to the influence of biological drives such as obedience, conformity, the desire to preserve self-esteem, which are often manifest as emotions and feelings (p. 190).

Damasio describes the “availability error”, and how participants tend to report that flying in a plane is more dangerous than driving a car. This is defective reasoning, because one is statistically more likely to be involved in a car crash than in a plane crash. Damasio suggests that this decision may be influenced by a “body-based negative influence” which does not fit with the statistics, but is likely survival-oriented, as more people survive car crashes than plane crashes. This conceptualization provides a critical role for emotions, which supports other conceptualizations of the rational role of emotions in goal achievement (Johnson-Laird & Oatley, 1992).

If System 1 heuristic processing (Stanovich, 1999) houses emotion, then Damasio's discussion provides a suitable conceptualization for the framing effect. Those who display the framing effect are opting for the riskless option in the lives saved version and the risky option in the lives lost version. When people are asked to gamble in the context of lives saves and lives lost, they may
be using a survival oriented heuristic. That is, they may choose the absolute value in the lives saved version so that they do not risk losing any lives in the risky version. In the lives lost version, they may choose the probability because there is a chance that less lives will be lost than with the absolute number option. This is a possible construal of the framing effect, and it would be consistent with the survival-oriented goals of System 1 processing (Damasio, 1994). The framing effect may be an instance in which the individual’s personal goals are in line with evolutionary goals.

The vehicles of cognitive ability and thinking dispositions were found to be associated with analytic response choices in the formal reasoning and decision-making problems. This is consistent with Stanovich’s (1999) prediction that both ability and dispositions would help to override heuristic, less optimal responses to satisfy the goals of disjunctive thinking. While System 1 or heuristic processes may be efficacious in many situations, the results here suggest that the tendency to contextualize and to rely on pragmatic cues is not efficacious in the domain of disjunctive thinking. Baron’s (1991) important distinction between cognitive abilities and thinking dispositions was also substantiated in the present study, as both sets of individual differences predicted unique variance on reasoning performance, in particular, on the formal reasoning tasks. Stanovich (1999) has implicated thinking dispositions which are related to epistemic regulation in argument evaluation (Stanovich & West, 1997) and belief bias (Sá, West, & Stanovich,1999). The present study implicates a new dispositional measure that is related to disjunctive reasoning performance: response regulation or the style of reflectivity/impulsivity, as measured by the MFFT (Kagan, Rosman, Day, Albert, & Phillips, 1964). The empirical demonstration of relationships between the MFFT and reasoning performance extends and
elaborates Stanovich’s (1999) original conceptualization of thinking
dispositions which have been primarily studied in belief bias (Sá, West, &
Stanovich, 1999) and argument evaluation (Stanovich & West, 1997).

Finally, the separate variance associated with abilities and dispositions in
the present study provides implications for remediating reasoning
performance in the domain of disjunctive thinking. In fact, in some cases,
thinking dispositions accounted for as much or more variance than what was
accounted for by cognitive ability. In the generically defined domain of
disjunctive thinking, or traveling down the branches of the decision tree
(Shafir, 1994), the evidence suggests that simple reflectivity, persistence, and
consideration of the alternatives can be as efficacious as high cognitive ability.
The results of this study support a Meliorist position of human rationality --
that is, “the Meliorist thinks that sometimes people are not reasoning very
well and that they could do much better.” (Stanovich, 1999, p. 7)

Limitations of the Present Study

Typical of many research studies in cognitive science, the
generalizability of the present study may be criticized due to a restrictive
selection of the participants. All of the participants in this study had attended
university, and one consequence of this sample is the restricted range in
cognitive ability across the sample. The average pro-rated full scale IQ score of
the population is 100, as calculated from the administered WAIS-R short form
according to Sattler’s (1992) formulas. In this sample, only 13.6% (n=17) of
participants had a pro-rated IQ score of less than 100, and 86.4% (n=108) of
participants had a pro-rated IQ score of 100 or greater. One main purpose of
this study was to explore the variance explained by stylistic tendencies, as
what the thinking dispositions were purported to measure. Consequently, the
separability of cognitive ability from these stylistic tendencies was statistically
more difficult to obtain, resulting in a possible underestimation of the role of stylistic aspects on disjunctive thinking performance.

The purpose of the informal tasks in this study was to extend the study of disjunctive thinking to more everyday domains of thinking. The measures that were taken as an index of disjunctive thinking included feature generation in the Consumer Task and my-side and other-side argument generation in the Argument Generation Task. Few results were obtained with these tasks in the current study, in particular, with respect to the individual difference measures. The prior experience and prior belief measures which were examined for each of these tasks indicated that the items and issues selected for each of these tasks were very heterogeneous. While on the one hand, the goal was to have breadth in the selection of tasks and issues in order to be able to speak more broadly about the domain of disjunctive thinking, the broad set of items and issues that were selected for the informal tasks came at a cost. In particular, the heterogeneity of the items and issues made the statistical analyses even more difficult to interpret because of the possibility that participants' were engaging with the items and issues differently than they were asked to do in the task. Other researchers have simply studied one item or issue, and have made generalizations about reasoning performance based on a single item (for example, Baron, 1995; Perkins, Farady, & Bushey, 1991) -- this is also inadequate, as eventually our generalizations must be demonstrated across a broad set of items and issues. In future studies of informal reasoning, a more homogeneous group of items or issues should be selected for study, or many more items should be used so that the effects of prior experience or prior belief can be studied systematically in the domain of disjunctive thinking.

Many of the formal reasoning and decision-making tasks had only two
or three response options from which to choose. When there are two response options, there is a 50% probability that the response selected may have been due to guessing or chance, and when there are three response options, there is a 33% chance that the response selected may have been due to guessing or chance. On some of the formal reasoning problems, the frequency of correct responses was very low, such as the married problem (12.8% correct) and the green levels problem (8.8% correct). When so few participants provide the correct solution, it becomes more important statistically that the small subset of correct responders be the ‘truly disjunctive reasoners.’ In subsequent studies with these two problems from Levesque (1986; 1989), participants were asked to provide confidence ratings on their response selection. The response selection on the problem and the confidence ratings together provided a better index of who had reasoned disjunctively on the problem. For example, a participant who has selected the correct response but reports no confidence in their response was interpreted as less likely to have reasoned disjunctively than a participant who selected the correct response and reported high confidence in their choice. In future studies with problems that have only two or three response selections, additional metrics, such as confidence ratings, should be used in order to increase the interpretable meaning of the response selection.

Finally, the current study included many tasks and many analyses. From a methodological design standpoint, the many tasks may have been overwhelming to complete for at least some of the participants. This aspect of the design would likely have worked against the goals of this study. That is, participants may not have been willing to engage disjunctively when faced with too many opportunities to use such a comprehensive strategy. The many analyses that are a consequence of such a large scale study with many tasks
are also recognized as a limitation, as one clearly loses statistical power as the number of contrasts increase (Kirk, 1982). Importantly, the primary goal of this research project was to study together those formal problems that have typically been studied in isolation and in separate literatures (Evans, Over, & Manktelow, 1993), and to extend the relevance of the principles of disjunctive thinking to more everyday informal problems. Despite these complexities, the analytic strategy used in this study was useful to help discern the generality of a very pervasive aspect of thinking, that is, the consideration and evaluation of alternatives.

**Summary and Educational Implications**

**Domain Generality and the Normative Issue**

Some evidence was found for a domain general skill related to disjunctive thinking in the formal reasoning problems, however, few relationships were obtained between the formal reasoning and decision-making problems suggesting specificity of performance on many of tasks. Based on some discussions from philosophical (such as, Nozick, 1993) and research literatures (Stanovich, 1999), alternative task construals are possible explanations of performance on some of these problems, suggesting that pure disjunctive consideration does not guarantee arrival at the optimal solution. On the informal tasks, some evidence was found for a generational fluency tendency which was unrelated to the type of disjunctive strategy necessary for solving the formal problems.

In Shafir’s (1994) conceptualization of disjunctive thinking, he argued that “a critical feature of thinking and deciding under uncertainty is the need to consider possible states of the world and their potential consequences for our beliefs and actions” (p. 403). If one treats situations of uncertainty as disjunctions of possible states and travels through all of the branches of the
decision tree, one is more likely to avoid violation of Savage's sure-thing principle (Shafir, 1994). Compellingly, Shafir (1994) displayed and argued how disjunctive consideration is fundamental for optimal outcomes in situations of uncertainty. On the one hand, the results of the present investigation displayed a consistent pattern of how consideration and evaluation of the consequences of each alternative leads to optimal reasoning in problems characterized by disjunctive states. However, the lack of relations obtained between many of the problems used in this study bring up some issues that must be considered in terms of understanding the implications of the overall lack of relations. For example, disjunctive thinking does not always work in favour of selecting options supported by alternative task construals, for example, with the Prisoner's dilemma and Newcomb's problem. Then, does this mean that disjunctive consideration is not useful?

Overall, the results of the present investigation suggest that disjunctive thinking leads to optimal choice selections, at least in the formal problems (see Table 38, page 142). Following Shafir (1994), it was also found that participants have difficulty with solving disjunctive problems that require the consideration and evaluation of a set of alternatives. Overall, the results of the present study support the useful and practical implications of considering and evaluating alternatives. Disjunctive thinking would likely never be harmful or uninformative. In fact, fully traveling all of the branches of the decision tree is an effective means for avoiding commission of the fundamental computation bias (Stanovich, 1999; Stanovich & West, 2000) and nonconsequentialist decisions (Baron, 1994).

**Should We Teach Disjunctive Thinking?**

One logical question that follows from the current investigation is whether the results of this study support the teaching of disjunctive thinking
Disjunctive Thinking

to students. As stated at the outset of the Introduction, the notion of considering the alternatives in a problem is a common way of describing an important part of good thinking (Nickerson, Perkins, & Smith, 1985; Bransford & Stein, 1993; Hayes, 1989; Beyer, 1988), and alternative consideration has been described as a key aspect in programs that claim to develop better thinking skills, such as Edward de Bono's Lateral Thinking. The results of the current study support the idea that fully traveling all of the branches of the decision tree will typically elucidate the optimal choice. In addition, the individual difference investigations provided further ideas about how to get individuals on route to thinking disjunctively. In particular, it was found that both cognitive ability and thinking dispositions predicted performance on the formal problems, and thinking dispositions remained an important predictor after cognitive ability was controlled. The focal dispositions were the Need for Cognition Scale (Cacioppo et al., 1996) and the Matching Familiar Figures Test (MFFT; Kagan, Rosman, Day, Albert, & Phillips, 1964). The findings with the Need for Cognition Scale and the MFFT reinforce the importance of persistence, engagement, intellectual curiosity, and reflectivity in disjunctive thinking performance. In particular, the idea of 'response regulation,' as measured by the total number of errors made on the MFFT, is empirically demonstrated to be a critical aspect in the domain of disjunctive thinking. These particular findings with the thinking dispositions measures also help to further conceptualize the vehicles to good reasoning, namely the role of thinking dispositions in System 2 processes (Stanovich, 1999; Stanovich & West, 2000). The implications of these findings are that reasoning performance in the domain of disjunctive thinking can be remediated or changed with some simple interventions, including inhibiting the first response, persisting longer, and teaching the value of searching and
evaluating alternatives as a means of achieving personal goals. The educational and practical implications of these findings are useful -- namely, these tendencies should not only be encouraged in educational programs and in practical situations, but they should be trained with the expectation that individuals are capable of using and implementing these propensities successfully.
References

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. Psychological Bulletin, 121, 219-245.


Dennett, D. C. (1978). *Brainstorms: Philosophical essays on mind and...*
Disjunctive Thinking

psychology. Cambridge, MA: MIT Press.


Newell, A. (1982). The knowledge level. *Artificial Intelligence, 18*, 87-


Please answer the following questions:

1) Have you ever purchased a breadmaker or helped someone select a breadmaker to purchase before?

   Yes   No   If Yes, estimate how many times in total: 

2) Have you ever purchased a t-shirt or helped someone choose a t-shirt to purchase before?

   Yes   No   If Yes, estimate how many times in total: 

3) Have you ever purchased a car or helped someone choose a car to purchase before?

   Yes   No   If Yes, estimate how many times in total: 

When buying different products, there are certain features that you take into consideration before buying the product. For example, if you buy a stereo, you will look for certain characteristics that should be considered, such as size of speakers, before buying your stereo. You will be presented with a list of products, and your task is to list those characteristics that you would take into consideration before purchasing the product.

1) A breadmaker
   a) ____________________________ b) ____________________________
   c) ____________________________ d) ____________________________
   e) ____________________________ e) ____________________________
   g) ____________________________ h) ____________________________
   i) ____________________________ j) ____________________________
   k) ____________________________ l) ____________________________
   m) ____________________________ n) ____________________________

2) A t-shirt
   a) ____________________________ b) ____________________________
   c) ____________________________ d) ____________________________
   e) ____________________________ e) ____________________________
   g) ____________________________ f) ____________________________
   i) ____________________________ h) ____________________________
   j) ____________________________ i) ____________________________
   k) ____________________________ j) ____________________________
   m) ____________________________ l) ____________________________
   n) ____________________________

3) A car
   a) ____________________________ b) ____________________________
   c) ____________________________ c) ____________________________
   e) ____________________________ d) ____________________________
   g) ____________________________ f) ____________________________
   i) ____________________________ h) ____________________________
   j) ____________________________ i) ____________________________
   k) ____________________________ j) ____________________________
   m) ____________________________ l) ____________________________
   n) ____________________________
Scoring Guide for Consumer Reports Task

NOTE: This scoring guide is based on pilot data from 12 participants. The frequency of responses given is indicated by the number in the parentheses.

1. Breadmaker

SIZE OF LOAF*²
-capacity (1)
-what amount of bread does it produce? (3)
-loaf size (5)
-quantity of bread made (size of loaf) (6)
-amount of bread produced per operation (10)

ENERGY EFFICIENCY
-energy efficiency (1)

EASE OF USE
-ease of use (buttons, functional, simplicity) (1,6)
-usability of the breadmaker (7)
-functional: easy to use (9)
-complexity of operation (10)
-ease of use (12)

APPEARANCE
-appearance (form, colour) (1)

COST*
-cost, price (1,3,4,5,7,8,10,11,12)
-cost at purchase time (6)
-price: affordable? (9)

QUALITY
-quality (material) (1)
-product quality (2)

OPTIONAL FEATURES
-actual functions, features offered (1)
-does it have a preset timer? (3)
-can I use a timer so that I can have fresh bread when I wake up in the morning? (4)
-does it have the features I wanted? (7)
-options: bake a cake? (9)
-other special features of the product (10)
-functions: basic vs. specialised functions (11)

² Consumer Reports indicated feature
REPUTATION
-reputation (Consumer Reports) (1)
-is it a reliable brand (7)

WARRANTY
-how long is warranty? (3)
-ease of warranty (6)

EASE OF CLEANUP*
-is it easy to clean? (3)
-ease of cleaning (6)

DURABLE
-is it a durable appliance? (3)
-durability and quality of merchandise (12)

LENGTH OF TIME
-how long does it take? (3)
-how quickly does it bake bread (as opposed to the oven) (5)
-time to produce the bread (10)

HEAVY
-is the appliance heavy (to move)? (3)

INTENSIVITY OF PROCESS
-is it a completely automatic process? (4)

SPACE IT OCCUPIES
-size (how much counter or storage space is available) (5)
-space it occupies (9)
-size (10)
-size: will it fit on my kitchen counter? (11)

RETRIEVING BREAD
-relative ease of retrieving baked loaf (5)

PRODUCT NAME
-brand name (5)
-brand of the breadmaker (6)
-the make (8)
-name brand (11)

AVAILABILITY OF SERVICE
-ease of service (6)

COST TO USE
-cost of use (6)
INSTRUCTIONS

-good instructions (12)

2. T-shirt

STYLE

-is it long or short sleeve? (3)
-what=s the style? (v-neck, crew neck) (3)
-style (is this acceptable for my taste in fashion) (5)
-cut/design (5)
-design: plain (no messages), round neck not polo (9)
-style (8,10)

FABRIC: TYPE AND WEIGHT*

-textile quality (1)
-what material is it made of? (cotton) (3)
-the fibre/fabric it is made up (4)
-fabric type and durability (5)
-quality of the material (i.e. is it 100% cotton?) (7)
-fabric (thick) (9)
-fabric (10)
-quality of material (12)

APPEARANCE

-appearance (texture) (1)
-if you really like it (2)
-aesthetic appeal (6)
-look and feel of the T-shirt - does it suit the wearer? (7)
-do I like it? (11)

SIZING

-size (1)
-what=s the size? (3)
-fit (6)
-fit: the right size (9)
-size availability (10)
-does it fit? (11)

COST*

-cost, price (1,3,6,7,8,10,12)
-how expensive it is (2)
-price: brand name prices (9)
-how much does it cost? (11)

COLOUR

-appearance (colour) (1)
-what=s the colour? (3)
-colour (5,7,8)
SHRINKAGE *
-preshrunk (1)
-whether it will shrink when I wash it (4)
-fabric (no shrinkage after washing) (9)

PRODUCT QUALITY
-product quality (2)
-quality (6,10)
-quality of workmanship (12)

WASHING INSTRUCTIONS
-how to wash it? (dry clean, or cold water) (3)
-washing instructions (5,6)
-wash instructions: easy to maintain (9)
-whether it is washable vs. dry cleaned (10)

COLOUR RETENTION *
-whether the colour will fade (4)

PRACTICALITY
-will it match my other clothes? (4)

COMFORT
-comfort (6)

WHERE PRODUCT WAS MADE
-country where it was made (7)
-where was it manufactured/made (12)

BRAND NAME
-brand name/designer label (10)

3. Car

STYLE/MODEL
-convertible: a little airing (9)
-style/marque (9)
-what model? (3)
-model (10)

2 DOOR VS. 4 DOOR
-type of car: specific or 2 door sedan (11)

COLOUR/APPEARANCE
-appearance (1)
-whether it appeals to me (6)
-colour (10,11)
COST
-cost, price (1,5,8,10,12)
-can handle with payments (2)
-what=s the price? (3)
-cost at purchase (6)
-overall cost (7)
-price: affordable (good deal) (9)
-price range (11)

OVERALL FUEL ECONOMY (*)
-fuel efficiency (1,6)
-what=s the gas mileage? (3)
-highly fuel efficient (4)
-mileage (5)
-mileage/consumption (economic fuel intake) (7)
-easy on gas (9)

SEATING CAPACITY
-capacity (1)
-how many people does it seat comfortably? (3)

REPUTATION
-reputation (e.g. market analysis, Lemon Aid, etc.) (1)
-name (dependability) (5)
-reputation of the vehicle (i.e. surveys/quality tests made on the car) (7)

EASE OF CARE
-difficult care (2)

EASE OF USE
-ease of use (1)
-ease of operation (6)

RIDE COMFORT (*)
-comfort (1)
-ride comfort (6)

BRAND
-what make? (3)
-brand (6)
-make (8)

WARRANTY
-what=s the warranty period? (3)

INTERIOR CONVENIENCE *
-ergonomic design of seats, console, steering wheel (12)
-convenience of gadgets and buttons (6)
TEST DRIVE
-test drive (10)

SAFETY INSPECTION
-is it safe/has it passed safety inspection? (4)
-safety record of car (11)

COST OF MAINTENANCE
-is it economical to maintain? (4)
-cost of parts and service (6)
-cost to operate (6)
-minimum maintenance (9)
-maintenance record of car (11)
-other monthly costs associated with it (11)
-projected cost of maintenance/repairs (12)

COST TO INSURE
-how much will insurance be? (4)
-cost to insure (6)

PREDICTED RELIABILITY *
-dependability involving safety stats, car life average (5)
-reliability of product (based on consumer reports) (12)

SIZE OF ENGINE
-size, 3 cylinder 6 cylinder (who will be using the car)

AGE OF CAR/USED
-new/old car? How many kilometres are on it previously (if any) (7)

OPTIONS/AESTHETIC
-luxuries (adjustable seats, coffee holder, stereo, thermometer, air bags) (5)
-options available (6)
-features (optional vs. standard) (10)
-does it have air-conditioning? (3)
-features and options of the car (11)

BRAKING SYSTEM
-does it include all features wanted (i.e. power steering, anti-lock braking) (7)

TRANSMISSION
-standard or automatic (5)

CRASH SAFETY
-crash worthiness (6)

PERFORMANCE*
-handling (6)
-manoeuvrability (6)

SIZE OF VEHICLE
-vehicle size (6)

STORAGE SPACE
-storage space (6)

SERVICE AVAILABILITY
-ease and availability of service (6)

REPUTATION OF DEALER
-dealership (8)
A) Think through the following issue carefully and feel free to take your time:

The real cost of a university education is $12000/year. Students are currently paying approximately $3500 in tuition. The difference is paid for by the taxpayer. University students should pay for the full cost of their university education.

Please write down arguments both for and against this position. Try to write as much as you can, and remember to try and give reasons both for and reasons against your position.

Reason #1: ________________________________________________________________

Reason #2: ________________________________________________________________

Reason #3: ________________________________________________________________

Reason #4: ________________________________________________________________

Reason #5: ________________________________________________________________

Reason #6: ________________________________________________________________

Reason #7: ________________________________________________________________
B) Think through the following issue carefully and feel free to take your time:

**People should be allowed to sell their organs.**

Please write down arguments both for and against this position. Try to write as much as you can, and remember to try and give reasons both for and reasons against your position.

Reason #1: ________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Reason #2: ________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Reason #3: ________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Reason #4: ________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Reason #5: ________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Reason #6: ________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Reason #7: ________________________________________________________________
________________________________________________________________________
________________________________________________________________________
C) Think through the following issue carefully and feel free to take your time:

**The cost of gasoline should be doubled to discourage people from driving.**

Please write down arguments both for and against this position. Try to write as much as you can, and remember to try and **give reasons both for and reasons against** your position.

Reason #1: __________________________________________________________

______________________________________________________________

Reason #2: __________________________________________________________

______________________________________________________________

Reason #3: __________________________________________________________

______________________________________________________________

Reason #4: __________________________________________________________

______________________________________________________________

Reason #5: __________________________________________________________

______________________________________________________________

Reason #6: __________________________________________________________

______________________________________________________________

Reason #7: __________________________________________________________

______________________________________________________________
Coding of Charles in Scotland Formal Reasoning Problem

All of the different response variations have been included in this Appendix to display how each has been coded for the statistical analyses.

1. Correct
- Charles is in Scotland OR June is in Wales and Kate is in Ireland.
- Either (June in Wales, Kate in Ireland, Charles not in Scotland) or (June not in Wales, Kate not in Ireland, Charles in Scotland)
- Either Charles is in Scotland OR June is in Wales and Kate is in Ireland.
- If Kate is in Ireland then June is in Wales. If Charles is in Scotland both Kate is not in Ireland & June is not in Wales.
- June is in Wales & Kate is in Ireland or Charles is in Scotland.
- If Kate is in Ireland then June is in Wales, and Charles is not in Scotland. If Kate is not in Ireland, then June is not in Wales, and Charles is in Scotland.
- If Charles is in Scotland, June is not in Wales, nor is Kate in Ireland. If he is not in Scotland, the 2 women are where the statement says they are.
- When Kate is in Ireland and June is in Wales, Charles is not in Scotland.
- When Kate is not in Ireland & June is not in Wales, Charles is in Scotland.
- If Charles is in Scotland, Kate isn’t in Ireland & June isn’t in Wales. If June is in Wales, Kate is in Ireland and vice versa.
- If Charles is not in Scotland, then June and Kate are in their respective places. If Charles is in Scotland, then June and Kate are not in their respective places.

2. June/Kate
- Where is Charles? June is in Wales/Kate is in Ireland.
- If June is in Wales/Kate is in Ireland.
- Kate could be in Ireland the same time June is in Wales.
- June and Kate can be in Wales and Ireland simultaneously.
- If June is in Wales, Kate is in Ireland, and if June is not, neither is Kate.
- June is not in Wales, Kate not in Ireland.

3. Partial Contingency
- If Charles is in Scotland, then June is not in Wales, nor is Kate in Ireland.
- If June is in Wales, and Kate is in Ireland, Charles is not in Scotland.
- Nothing is for sure except that if Charles is not in Scotland, then June is in Wales (for sure) and Kate is in Ireland.
- If Charles is not in Scotland, June is in Wales and Kate in Ireland. If June is in Wales, Charles is not in Scotland and Kate is in Ireland.
- Nothing follows. If Charles is in Scotland then June isn’t in Wales and Kate isn’t in Ireland. But we do not have enough information to say Charles is in Scotland.
- Charles is never in Scotland with a woman, & when Charles is in Scotland, June and Kate are not in Wales or Ireland (respectively).
4. Nothing
- nothing
- Nothing all are mutually exclusive.
- Nothing; can’t say for sure who is where.

5. Blank
- (space left blank)
- Cannot be determined
- Don’t know

6. Charles is in Scotland
- Charles is in Scotland.
- Charles is in Scotland?
- That Charles should be in Scotland.
- It seems likely Charles is in Scotland.
- Charles is most likely in Scotland.
- That Charles is in Scotland (he’s got the greater odds).
- 50% chance Charles is in Scotland.
- There’s a good chance that Charles is in Scotland (about 50%)
- Charles is in Scotland -> most likely of all two facts, although it is not 100%.
- Charles is in Scotland 50% of the time.
- Charles may or may not be in Scotland.
- Charles is more likely to be in Scotland than the other two in their place.
- Charles is in Scotland only. We don’t know where June or Kate is.
- Charles is in Scotland but June and Kate are not.

7. Other Incorrect
- There are people living in the United Kingdom (where? I am not sure)
- June, Charles, and Kate are somewhere in the UK or Ireland.
- These names are typically used in these countries but in others no.
- If June is not in Wales, Kate cannot be in Ireland and vice-versa.
- At least one of these people is on the British Isles.
- Kate is in Ireland.
- This question was made by someone from the United Kingdom.
- 1 or 2 people are in Great Britain, but not all three. If it is only 1 person then that person is Charles.
- Charles is in Scotland and June is in Wales.
- Is Charles having an affair?
- Anyone could be anywhere...very inconclusive.
- Can June in Wales and Kate in Ireland exist at the same time?
- Charles is not in Wales or Ireland.
- Someone is in Great Britain.
- (June in Wales and/or Kate in Ireland) or Charles in Scotland. All 3 could be in the same place, or Charles could be with either of the other 2; I won’t speculate on their relation.
- June & Kate can be in Wales & Ireland respectively only when Charles is in Scotland, and vice-versa.
7. Other Incorrect (continued)
-June is in Wales.
-June & Charles are in Great Britain while Kate is in Ireland.
-Charles is in Ireland or June is in Scotland, but not both.
-That we are not sure exactly where these 3 are.
-June is in England, Charles is in Scotland, Kate is in Ireland.
-June is in Wales or Kate is in Ireland, but not both.
-Somebody is in the UK, either just Charles or both June and Kate.
-Kate is in Ireland or _____, but not both.
-If June is in Wales then Charles is in Scotland.
-Charles has a better chance of being somewhere.
-Kate is definitely not in Wales or Scotland, June is definitely not in Ireland or Scotland, & Charles is definitely not in Ireland or Wales.
-Somebody is somewhere, but who and where can't be determined.
Sample Item from the Matching Familiar Figures Test (MFPT)

**Target Item**

**Six Possible Options**
Scales Used In The Thinking Dispositions Questionnaire

A. Need for Cognition Composite

Need for Cognition Scale (Cacioppo et al., 1996)

1. I would prefer complex to simple problems.
2. I like to have the responsibility of handling a situation that requires a lot of thinking.
3. Thinking is not my idea of fun. (Reverse Scored)
4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities. (Reverse Scored)
5. I try to anticipate and avoid situations where there is a likely chance I will have to think in depth about something. (Reverse Scored)
6. I find satisfaction in deliberating hard and for long hours.
7. I only think as hard as I have to. (Reverse Scored)
8. I prefer to think about small, daily projects to long-term ones. (Reverse Scored)
9. I like tasks that require little thought once I've learned them. (Reverse Scored)
10. The idea of relying on thought to make my way to the top appeals to me.
11. I really enjoy a task that involves coming up with new solutions to problems.
12. Learning new ways to think doesn't excite me very much. (Reverse Scored)
13. I prefer my life to be filled with puzzles that I must solve.
14. The notion of thinking abstractly is appealing to me.
15. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.
16. I feel relief rather than satisfaction after completing a task that required a lot of mental effort. (Reverse Scored)
17. It's enough for me that something gets the job done; I don't care how or why it works. (Reverse Scored)
18. I usually end up deliberating about issues even when they do not affect me personally.

Subset Of Items Used From The Typical Intellectual Scale (Goff & Ackerman, 1992)

1. I don't let my ideas run my life; I wouldn't go out of my way to write them down or tell others about them. (Reverse Scored)
2. I feel most intellectually able when I'm deeply involved in a problem.
3. I enjoy thinking out complicated problems.
4. There are very few topics that bore me.
5. I am philosophically inclined, that is, inclined to philosophize about things.
6. Sometimes I like to consider concepts even if they may be of no practical consequence.
7. At times I have been so entertained by the cleverness of a crook that I have hoped he would get away with it.
8. I am more interested in athletics than intellectual things. (Reverse Scored)
9. I have only one or two real hobbies. (Reverse Scored)
10. I must admit I have no great desire to learn new things.
11. I am not really bothered by learning something incompletely. (Reverse Scored)
12. Ignorance is bliss. (Reverse Scored)
13. I try to set aside definite periods of time for thinking about things that interest me.
14. I have difficulty thinking in new and unfamiliar situations. (Reverse Scored)
15. I am an intellectual.

B. Actively Open-Minded Thinking Composite

Actively Open-Minded Thinking Scale (Stanovich & West, 1997)

1. Changing your mind is a sign of weakness. (Reverse Scored)
2. A person should always consider new possibilities.
3. Intuition is the best guide in making decisions. (Reverse Scored)
4. If I think longer about a problem I will be more likely to solve it.
5. Basically, I know everything I need to know about the important things in life. (Reverse Scored)
6. Considering too many different opinions often leads to bad decisions. (Reverse Scored)
7. People should always take into consideration evidence that goes against their beliefs.
8. Difficulties can usually be overcome by thinking about the problem, rather than through waiting for good fortune.
9. There is nothing wrong with being undecided about many issues.
10. Coming to decisions quickly is a sign of wisdom. (Reverse Scored)

Openness-Ideas From The Revised NEO Personality Inventory (Costa & McCrae, 1992)

1. I often enjoy playing with theories or abstract ideas.
2. I find philosophical arguments boring. (Reverse Scored)
3. I enjoy solving problems or puzzles.
4. I sometimes lose interest when people talk about very abstract, theoretical matters. (Reverse Scored)
5. I enjoy working on "mind-twister"-type puzzles.
6. I have little interest in speculating on the nature of the universe or the human condition. (Reverse Scored)
7. I have a lot of intellectual curiosity.
8. I have a wide range of intellectual interests.
Dogmatism Subscale (Rokeach, 1960; and two items from Paulhus & Reid, 1991)

1. Of all the different philosophies which exist in the world there is probably only one which is correct.
2. Even though freedom of speech for all groups is a worthwhile goal, it is unfortunately necessary to restrict the freedom of certain political groups.
3. There are two kinds of people in this world: those who are for the truth and those who are against the truth.
4. Often, when people criticize me, they don't have their facts straight. (Paulhus & Reid, 1991)
5. No one can talk me out of something I know is right. (Paulhus & Reid, 1991)
6. A group which tolerates too much difference of opinion among its members cannot exist for long.
7. There are a number of people I have come to hate because of the things they stand for.
8. My blood boils over whenever a person stubbornly refuses to admit he's wrong.
9. Most people just don't know what's good for them.

Categorical Thinking From The Constructive Thinking Inventory (Epstein & Meier, 1989)

1. There are basically two kinds of people in this world, good and bad.
2. I think there are many wrong ways, but only one right way, to almost anything.
3. I tend to classify people as either for me or against me.

C. Other Dispositions

Objectivism Scale (Leary et al., 1986)

1. I seek as much information as possible before making decisions.
2. I think the answers to most questions in life can be found through careful, objective analysis of the situation.
3. I do not like to be too objective in the way I look at things. (Reverse Scored)
4. Trying to be highly objective and rational does not improve my ability to make good decisions. (Reverse Scored)
5. I see myself as a rational and objective person.
6. After I make a decision, it is often difficult for me to give logical reasons for it. (Reverse Scored)
7. I gather as much information as possible before making decisions.
8. The solution to many problems in life can not be found through an intellectual examination of the facts. (Reverse Scored)
9. I try to employ a cool-headed, objective approach when making decisions about my life.
10. I am only confident of decisions that are made after careful analysis of all available information.
11. I tend not to be particularly objective or logical in my approach to life. (Reverse Scored)

Vigilance Subscale From The Melbourne Decision Making Questionnaire (Mann et al., 1997)
1. I like to consider all of the alternatives.
2. I try to find out the disadvantages of all alternatives.
3. I consider how best to carry out the decision.
4. When making decisions I like to collect lots of information.
5. I try to be clear about my objectives before choosing.
6. I take a lot of care before choosing.

Openness-Values From The Revised NEO Personality Inventory (Costa & McCrae, 1992)
1. I believe letting students hear controversial speakers can only confuse and mislead them. (Reverse Scored)
2. I believe that laws and social policies should change to reflect the needs of a changing world.
3. I believe we should look to our religious authorities for decisions on moral issues. (Reverse Scored)
4. I believe that the different ideas of right and wrong that people in other societies have may be valid for them.
5. I believe that loyalty to one's ideals and principles is more important than "open-mindedness." (Reverse Scored)
6. I consider myself broad-minded and tolerant of other people's lifestyles.
7. I think that if people don't know what they believe in by the time they're 25, there's something wrong with them. (Reverse Scored)
8. I believe that the "new morality" of permissiveness is no morality at all. (Reverse Scored)

Belief in Prayer (Paulhus & Reid, 1991)
1. In one way or another, God answers all my prayers.

Paranormal Beliefs Subscale
1. It is advisable to consult your horoscope daily. (belief in astrology; from Jones, Russell, & Nickel)
2. Astrology can be useful in making personality judgments. (belief in astrology; from Jones, Russell, & Nickel)
3. I have personal possessions that bring me luck at times. (concept of
Disjunctive Thinking
Appendix F

luck; from Tobacyk & Milford, 1983)
4. The number 13 is unlucky. (concept of luck; from Tobacyk & Milford, 1983)
5. It is bad luck to have a black cat cross your path. (concept of luck; from Tobacyk & Milford, 1983)
6. Opening an umbrella indoors will increase one's chances of misfortune in the near future. (concept of luck; from Tobacyk & Milford, 1983)

Impression Management (subset of items taken from Paulhus, 1991)

1. I always obey laws, even if I'm unlikely to get caught.
2. I don't gossip about other people's business.
3. I sometimes tell lies if I have to. (Reverse Scored)
4. There have been occasions when I have taken advantage of someone. (Reverse Scored)
5. I have said something bad about a friend behind his or her back. (Reverse Scored)
Appendix G

Vocabulary Checklist

Below you will see a list of 60 letter strings. Some of the strings are actual words and some are not. You are to read through the list of items and indicate whether or not you think the letter string is a word by putting a check mark next to those that you know to be words. Do not guess, but only check those who you know to be words.

1. absolution ____
2. arrate ____
3. asinine ____
4. audible ____
5. ceiloplaty ____
6. clandestine ____
7. comectial ____
8. concurrent ____
9. confluence ____
10. connote ____
11. denotation ____
12. denouement ____
13. disconcert ____
14. disler ____
15. dropant ____
16. epicurean ____
17. eventuate ____
18. fisigenic ____
19. gustation ____
20. heuristic ____
21. hyplexion ____
22. ineffity ____
23. inflect ____
24. inundate ____
25. irksome ____
26. lacuna ____
27. langour ____
28. laudatory ____
29. litany ____
30. metenetion ____
31. neotatin ____
32. niche ____
33. nonquasity ____
34. nuance ____
35. nitrous ____
36. optimize ____
37. plabbage ____
38. polarity ____
39. potomite ____
40. purview ____
41. recidivism ____
42. reportage ____
43. reverent ____
44. rochead ____
45. selement ____
46. sheal ____
47. sparkhouse ____
48. stratagem ____
49. subjugate ____
50. substratum ____
51. suffuse ____
52. tenacious ____
53. tradured ____
54. tuncier ____
55. ubiquitous ____
56. unction ____
57. unnanal ____
58. wanderlust ____
59. waterfowl ____
60. xenophobia ____
## Order of Task Administration

1. Informed Consent
2. Demographics Form
4. Prior Purchases Questionnaire for The Consumer Task [Booklet #1] *(See Method section starting on page 39 for task description and Appendix A on page 173 for actual task used)*
5. Dispositions Questionnaire [Booklet #1] *(See Method section starting on page 52 for description of task and Appendix F starting on page 189 for scales used)*
6. Argument Generation Rating Task [Booklet #1] *(See Method section starting on page 46 for description of task)*
7. WAIS-R Vocabulary Subtest *(See Method section starting on page 55)*
8. WAIS-R Block Design Subtest *(See Method section starting on page 55)*
10. Argument Generation Task [Booklet #2] *(See Appendix C starting on page 182)*
11. Break – Optional
13. Vocabulary Checklist [Booklet #3] *(See Appendix G on page 194)*
15. Matching Familiar Figures Test (MFFT) *(See Appendix E on page 188 for sample item)*
16. Raven's Progressive Matrices *(See Method section starting on page 55)*