INDIVIDUAL DIFFERENCE AND DEBIASING STRATEGIES IN THINKING BIASES AND ATTITUDE POLARIZATION

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

---

by

Shih-Ching Wang

August, 2013

Examining Committee Members:

Dr. Eric M. Eisenstein, Advisory Chair, Marketing and Supply Chain Management

Dr. Anthony Di Benedetto, Marketing and Supply Chain Management

Dr. Nathan Fong, Marketing and Supply Chain Management

Dr. Dipankar Chakravarti, External Member, Johns Hopkins University

ABSTRACT

Most consequential decisions are made by more than two people. People frequently argue with each other to make better decisions. However, most decision making research usually only involves small purchases and individual decisions. The lacking of investigation in high-cognition decisions and argumentative settings is the motivation of this research.

Researchers studying decision making have largely focused on how the decisions that people make are affected by task characteristics, and how labile decisions are with respect to situational factors. However, the fact that many preferences are constructed does not imply that all constructions are equally good. That people differ from each other in thinking is obvious. How and why they differ is less clear.

Therefore, the first two studies are foundational studies in order to find out the most important and germane individual difference factor that may be the best predictor of thinking ability, including argument generation quality, evaluation ability, and debiasing ability. I found that logical reasoning ability is the best predictor of both thinking and debiasing ability.

Argumentative Theory (Mercier &Sperber, 2011) claims that when reasoning is used in argumentative contexts, the confirmation bias contributes to an efficient form of division of cognitive labor, and then lead to better decisions and attitude depolarization. In study 3, I provided implication evidence to show that either arguing with the other person or viewing arguments from the opposite perspective may lead to attitude depolarization. Most interestingly, individual differences did moderate the main effects.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without all the help from so many people in so many different ways. This page also reflects the relationships with many generous and inspiring people I have met since beginning my graduate work.

I would like to express the deepest appreciation to my advisor, Dr. Eric Eisenstein, who has the substance of a genius. His wisdom, knowledge and commitment to the highest standards inspired and motivated me.

To my committee members Anthony Di Benedetto, Dipankar Charkravati and Nathan Fong for their encouraging words, thoughtful comments, time, and attention.

To my colleagues for accompanying me through these years: Dale Cake, David DiRusso, Ebru Genc, Ashley Goerke, Le Ki, Wooyang Kim, Guihan Ko, Mark Lang, Suzanne Makarem, Mike Obal, Miranda Rahman, Ellen Thomas, Dan Young, and Dan Zhang.

To Administrative Specialist of Marketing Department for assisting me through all these years: Queene Mays.

To Fox Doctoral Program Assistant Director for assisting me with all the administrative tasks necessary for completing my doctoral program: Lisa Fitch.

To my parents, siblings, aunts, all the relatives and friends for their unconditional love, support and understanding during the long years of my PhD education.

And finally, to Leonard Teo, my boyfriend and cheerleader. Thank you for walking me through the toughest year and always being there for me.

TABLE OF CONTENTS

Page

ABSTRACT ii

ACKNOWLEDGMENTS iii

LIST OF TABLES viii

LIST OF FIGURES x

CHAPTER

1. INTRODUCTION 1

2. THEORETICAL BACKGROUND 7

Good Thinking 7

Critical Thinking 9

Confirmation Bias / Myside Bias 11

Reasoning and the Heuristics and Biases Debate 13

Individual Differences in Rational Thought and Biases 16

Cognitive Ability 16

Thinking Dispositions 19

Actively Open-minded Thinking 21

Need for Cognition 22

Dual Process Theories 23

Cognitive Reflection Test 25

Logical Reasoning Ability 27

3. STUDY ONE 30

Introduction 30

Conceptual Framework 32

Cognitive Ability 34

Thinking Dispositions 35

Logical Reasoning Ability 37
Cognitive Reflection Test 38
Methodology 39
Procedure 39
Measurements 40
Dependent Variables 40
Independent Variables 41
Results 44
Argument Quality 44
Myside Bias 47
Conclusion 50
Discussion 51
Other Interesting Findings and Future Study 52
4. STUDY TWO 55
Introduction 55
Conceptual Framework 58
Cognitive Ability 59
Thinking Dispositions 60
Logical Reasoning Ability 62
Cognitive Reflection Test 63
Methodology 65
Experimental Design 65
Pretest 66
Procedure 66
Measurements 68
Dependent Variables 68
Independent Variables 69

Results 72
Argument Quality 72
Argument Generation Task 72
Argument Evaluation Task 74
Myside Bias 76
Argument Generation Task 76
Argument Evaluation Task 77
Manipulation Check 79
Argument Order Effect 79
Conclusion 80
Other Interesting Findings and Future Study 80
5. STUDY THREE 86
Introduction 86
Literature Review 90
Confirmation Bias 90
Argumentative Theory 91
Attitude Polarization and Confidence Change 93
Methodology 95
Experimental Design 95
Procedure 96
Measurements 98
Dependent Variables 98
Independent Variables 98
Results 99
Attitude Polarization 99
Argument Generation Task 99
Argument Evaluation Task 101

Confidence Change 106
Argument Generation Task 106
Argument Evaluation Task 110
Conclusion 114
Discussion 116
GENERAL DISCUSSION 117
REFERENCES 123
APPENDICES
A. CONSENT FORM 147
B. NEED FOR COGNITION SCALE 149
C. ACTIVELY OPEN-MINDED THINKING SCALE 150
D. LOGICAL REASONING ABILITY QUESTIONS 153

LIST OF TABLES

Table Page

1. Possible Framework for Future Studies 34
2. SLS Analysis Predicting Initial Argument Generation Quality 46
3. Repeated Measures MANOVA Analysis Predicting Argument Generation Quality 46
4. SLS Analysis Predicting Initial Myside Bias 48
5. Repeated Measures MANOVA Analysis Predicting Myside Bias 48
6. SLS Analysis Predicting Initial Otherside Argument Counts 49
7. Repeated Measures MANOVA Analysis Predicting Otherside Argument Counts 49
8. Repeated Measures MANOVA Analysis Predicting Strong Opinion 54
9. Repeated Measures MANOVA Analysis Predicting Confidence Level 54
10. SLS Analysis Predicting Initial Argument Generation Quality 74
11. Repeated Measures MANOVA Analysis Predicting Argument Generation Quality 75
12. SLS Analysis Predicting Argument Evaluation Ability 75
13. SLS Analysis Predicting Initial Myside Bias in Argument Generation Task 77
14. Repeated Measures MANOVA Analysis Predicting Myside Bias in Argument Generation Task 78
15. SLS Analysis Predicting Myside Bias in Argument Evaluation Task 78
16. SLS Analysis Predicting Argument Order Effect 79
17. SLS Analysis Predicting Attitude Polarization in Argument Generation Task 83

18. SLS Analysis Predicting Confidence Change in Argument Generation Task.....83

19. SLS Analysis Predicting Attitude Polarization in Argument Evaluation Task 84

20. SLS Analysis Predicting Relevant Counts in Argument Generation Task 84

21. SLS Analysis Predicting Otherside Counts in Argument Generation Task.....85

LIST OF FIGURES

Figure Page

1. Study 1 Framework 33

2. Study 2 Framework in Both Argument Generation and Evaluation Tasks 65

3. Interaction Effect of Timing and Argument Generation Task on Attitude Polarization 100

4. Interaction Effect of Timing and Argument Evaluation Task on Attitude Polarization 102

5. Interaction Effect of Logical Reasoning Ability and Argument Evaluation Task on Attitude Polarization 103

6. Interaction Effect of Cognitive Reflection Test and Argument Evaluation Task on Attitude Polarization 105

7. Interaction Effect of Timing and Argument Generation Task on Confidence Change 107

8. Interaction Effect of Cognitive Reflection Test and Argument Generation Task on Confidence Change 109

9. Interaction Effect of Need for Cognition and Argument Generation Task on Confidence Change 110

10. Interaction Effect of Timing and Argument Evaluation Task on Confidence Change 111

11. Interaction Effect of Actively Open-minded Thinking and Argument Evaluation Task on Confidence Change 113

CHAPTER 1

INTRODUCTION

Most marketing research focuses on relatively small purchases (e.g., grocery) and these purchases usually do not require particularly high-cognition decision making processes. Consider the now familiar conclusion frequently stated in the introduction of behavioral decision theory (BDT) articles: "There is a growing consensus that preferences are typically constructed when decisions are made, rather than retrieved from a master list of preferences stored in memory. In particular, preferences are influenced by the method of preference elicitation, the description of the options and the choice context." This basic conclusion has been recognized as "one of the main themes that has emerged from behavioral decision research during the past three decades" (Lichtenstein & Slovic, 2006).

Researchers studying decision making have largely focused on how the decisions that people make are affected by task characteristics, how labile decisions are with respect to situational factors, and the degree to which preferences are constructed at the time of the decision. However, the fact that many preferences are constructed does not imply that all constructions are equally good. For example, if we accept that being "rational" means that "rational" thought helps people achieve their goals (Baron, 2000), then we conclude that, in almost all decision environments, good thinking involves sufficient search for possibilities and evidence, and fairness in the evaluation of the obtained evidence (e.g., Baron, 1985). Even early proponents of constructed preferences conceded that non-situational, individual-level characteristics could exert strong influences on both strategies for approaching a decision problem and the construction of

preferences. For example, Payne, Bettman, and Johnson (1993) note that the expertise of the decision maker could lead to substantial differences in approach and outcome, and that the other characteristics of the decision maker could also matter.

An issue that remains open is the degree to which people can develop expertise in the process of decision making itself. Many researchers conclude that expertise is strongly domain-bound, with little or no far-transfer (see Barnett & Ceci, 2002, for a review). Other researchers, however, argue that certain individual differences that could be classified as forms of "expertise in thinking" have the characteristics of traits, in that they are stable, not domain specific, and can affect entire substrates of thought (e.g., Keith E. Stanovich, 2000). In the first study, I support the view that logical reasoning ability predict fewer biases in reasoning and increased quality of generated arguments, including both-sides evidence, in constructed preference-like tasks across different domains. Most importantly, logical reasoning ability is not natural, and can be improved by training. Also, this trait can be carried over into different domains.

Consumers frequently make suboptimal decisions. In past research, this tendency has been attributed to a number of causes, including having difficulty in decoupling from prior opinions and beliefs (Stanovich & West, 2007), and using heuristics (Tversky & Kahneman 1974) while making decisions. Many of these characteristics become particularly salient when decisions require analysis of a complex set of alternatives (Arkes, 1991; Larrick, 2004). In particular, confirmation bias (also called myside bias), and one-sided bias are all likely to be exacerbated when decisions are complex. Moreover, increased decision complexity allows more opportunities for motivated reasoning, selective recruitment of reasons, and "polarization"-like effects. It has been

proposed that many of these problems can be explained by a simple idea: people do not generate arguments in order to "get the right answer", but rather to win others to their way of thinking (Mercier & Sperber, 2011; Nickerson, 1998).

Accepting the existence of a normative-descriptive gap raises the question of how the gap might be closed. Arkes (1991) puts biases into three categories: Strategy-Based Errors, Association-Based Errors, and Psychophysically-Based Errors. Confirmation bias/myside bias belongs to Association-Based Errors. The basic problem of association-based processes is an overly narrow sample of evidence. Therefore, I hope to use this series of studies to provide some evidence for association-based errors. There are two main approaches to debias in association-based errors.

One approach to prescription focuses on modifying the cognitive strategies of the individuals. Meliorists believe that everyday reasoning falls far short of the ideal, but can be improved through experience and education (Nisbett, 1993). Stanovich and his colleagues have provided some interesting evidence on this point in support of the Meliorists. They observe that there is always a subset of decision makers who give a normative response on a decision task, indicating that normative cognitive strategies are not unattainable, but are systematically held and used by some individuals. However, they couldn't replicate the experiments and find the strong correlation between cognitive ability, thinking dispositions and the debiasing ability all the time (Stanovich & West, 1997, 1998, 2000). Therefore, in the first study, I investigated the best predictor (among all the possibly potential individual differences, including innate and acquired ones) of good thinking and debiasing ability.

Despite the different emphases in these approaches, they share a common implication: debiasing requires intervention. However, there is debate about individual difference in debiasing strategies. In a series of studies, I demonstrated that the acquisition of certain habits of thinking and associated skills markedly improves the quality of reasoning among a diverse set of participants. The individual differences I identify are mutable, amenable to training, and are of a magnitude that equals or exceeds documented trait-based individual differences. In addition, the wide applicability of these results suggests that the identified individual differences constitute "meta-expertise", in the sense that, unlike most developed expertise that is strongly domain-bound and for which transfer to novel situations is nearly nonexistent, this type of expertise is readily applicable across diverse and substantially different domains. Finally, I also tested the simple but general cognitive strategy of "consider the opposite" (Arkes, 1991) in the first study, and showed that the individual differences moderate the effect of this individual debiasing strategy. I found that logical reasoning ability is the best predictor of argument generation quality and the ability to avoid confirmation/myside bias in the argument generation tasks.

In real life scenarios, people may have the opportunity to be the judge of one-sided or two-sided arguments, so having the ability to discriminate the good arguments from the bad ones is really important, especially when people need to make the judgments/decisions right after they see arguments from both sides. In Study 2, I showed respondents two sets of arguments, including four pro and four con arguments in each issue, and asked them to rate the persuasiveness of each argument. Also, I asked a series of questions about their ratings, attitude, and confidence. I would like to see how their

prior attitudes influence individuals' argument evaluation ability. Also, I discussed the moderating effect of individual differences. Most importantly, I found that confirmation bias is not the same in argument generation and argument evaluation tasks. On the one hand, some people may be able to avoid confirmation bias during memory retrieval stage, but not in information processing stage. On the other hand, some individuals are able to conquer the confirmation bias during information processing, but not in memory retrieval or information search stage.

The second approach to prescription is to expand possible strategies to include techniques external to the decision maker. Individual reasoning can approach normative standards through the use of tools (Larrick, 2004). Such techniques include using groups to replace individuals, improving information processing through decision aids and information displays, supplementing intuitive decision making with formal decision analysis, and replacing individual judgment entirely with statistical models. In Study 3, I compared some of the strategies that can be performed by human, such as "Considering the Opposite" (i.e., prompting individuals to consider the arguments from the opposite view), "Groupthink" (i.e., joining a group discussion having the same view as the individual), and "Argumentative/Debate" (i.e., joining a group discussion having the opposite view from the individual). Also, I found some moderating effects of individual difference on these strategies.

Further, consumers do not only make decisions by themselves. They often need to make purchase decisions with their family members, especially the important, complicated decisions. Also, group discussion is one of the debiasing strategies mentioned in the literature. Using groups to improve decisions ultimately depends on

assembling a group with diverse experience and training, and then following a process that preserves the diversity of perspectives. If run effectively, groups generate their own "consider-the-opposite" process. However, the outcomes are controversial, and according to the Argumentative Theory, when reasoning is used in argumentation setting, the confirmation bias contributes to an efficient form of division of cognitive labor (Mercier & Sperber, 2011). Therefore, in the third study, I examined group discussion strategy in different settings-group think and argumentation setting. I would like to see if group decision leads to attitude polarization or moderation/depolarization? I proposed that people who are high in logical reasoning ability, they will take both sides' arguments into account while they construct their preferences. Therefore, arguing with the other side will lead them to depolarization, compared to people low in logical reasoning ability.

Finally, the Argumentative Theory argues that the efficiency of 'confirmation bias' only works in argument generation task, but not in argument evaluation task; however, no empirical evidence has been provided to support their saying. Therefore, I also examined the effects of three different combinations (all consistent statements, all inconsistent statements, and mixed statements) in argument evaluation task. I found some supporting findings and some contradicting results to the Argumentative Theory. The results in Study 3 showed that people did perform better in argumentative settings, but the strategies not only worked in argument generation task, but also worked in argument evaluation task. Most interestingly, individual differences did moderate the main effects.

CHAPTER 2

THEORETICAL BACKGROUND

In this section, I review four streams of literature most germane to this research: good thinking, confirmation bias/myside bias, heuristic and biases debate, individual differences in rational thoughts and biases, and attitude polarization.

Good Thinking

Rational thinking, the ideal kind of thinking, means we think based on our own best interests in order to achieve our goals (Baron, 2000). Baron indicated that there are three basic types of thinking that we have to do for achieving our goals: thinking about decisions, beliefs, and our own goals. In most general sense, thinking is a method of looking for and choosing among potential possibilities, considering possible actions, beliefs and personal goals.

> The human understanding when it has once adopted an opinion draws all things else to support and agree with it. And though there be a greater number and weight of instances to be found on the other side, yet these it either neglects and despises, or else by some distinction sets aside and rejects, in order that by this great and pernicious predetermination the authority of its former conclusion may remain inviolate.
>
> Francis Bacon

Baron (1985, 1991) has been arguing that good thinking involves both optimal search for possibilities, evidence and goals, and also fairness in the evidence search and

in inference-making in purpose of maximizing the expected utility of the outcome in terms of the individual's goals. Search should be 'optimal' in terms of thinker's personal goals, including the goal of minimizing the thinking cost and efforts, and 'fair', meaning that it is not influenced by factors unrelated to the goals of the thinking itself.

Baron (2000) concludes that poor thinking could be attributed to three reasons. First, our search misses something that it should have been discovered, or we act with high confidence after little search. Second, we seek evidence and make inferences in ways that prevent us from choosing the best possibility. Third, we may overthink. He thinks that the second of these problems seems to be the most serious. People tend to look for evidence and make inferences to favor possibilities that already appealing to them, and to ignore the evidence against the possibilities they prefer.

Therefore, poor thinking could be caused by too little search, by overconfidence, and by biases in favor of possibilities that are favored initially. On the contrary, good thinking consists of (1) optimal thorough search, (2) appropriate confidence in the quality of thinking done, and (3) fairness to all the possible outcomes. Baron (2000) argues that good thinking is the one that advises and helps people to become more actively open-minded. It counteracts the major thinking biases (e.g., confirmation bias and myside bias), and allows consideration of new evidence, goals and possibilities (Baron, 1991).

Being actively open-minded is a good method to counteract the biases favoring prior beliefs. In general, good thinking is most likely to lead to true beliefs with appropriate confidence. The main advantage of true beliefs is that they allow us to make better decisions, and then are more likely to achieve our goals (Baron, 2000).

The most important part of good thinking is 'fairness to all possible outcomes'. However, people are inclined to favor the outcomes that they already prefer initially, both in searching for evidence and in making inferences. Also, they tend not to search for evidence against their favored possible outcomes or ignore it when they find it. Perkins names these two characteristics "confirmation bias" (Perkins, Bushey, & Faraday, 1986).

In many situations, experts seems not to be good thinkers because they know the answers to most questions, experts generally do not think that often as novices do. Although good thinking and expertise both contribute to success in achieving goals, we may not see them as the same thing (Baron, 1985). We should not presume that the understanding of expertise will solve the problem of providing standards for good thinking (Baron, 1991).

Critical Thinking

Critical thinking is part of good thinking (Facione, 2011). In the psychological and educational literature, the discussions of critical thinking point that it is imperative for critical thinkers have decontextualized reasoning styles, being able to decouple their prior beliefs and opinions from the evaluation of arguments and evidence (Baron, 1991, 1995, 2000; J. Evans, 2002; Galotti, 1989; Klaczynski, 1997; Klaczynski & Gordon, 1996; Klaczynski, Gordon, & Fauth, 1997; Kuhn, 1993; Norris & Ennis, 1989; Paul, 1984, 1987; Perkins, 1995; Stanovich, 1999, 2000, 2004; Stanovich & West, 2007; Sternberg, 1997, 2001, 2003; Voss, Perkins, & Segal, 1991).

Researchers have listed one characteristic of critical thinking as the disposition "to accept statements as true even when they don't agree with one's own position" (Zechmeister & Johnson, 1992), "to reason from starting points with which they disagree without letting the disagreement interfere with their reasoning" (Norris & Ennis, 1989), and "recognize the fallibility of one's own opinions, the probability of bias in those opinions and the danger of differentially weighting evidence according to personal preferences" (Nickerson, 1987). Critical thinking is self-guided, self-disciplined thinking that attempts to reason at the highest level of quality in a fair-minded way. People, who consistently think critically, attempt to live rationally, reasonably, and empathically. People who are high in critical thinking strive to diminish the power of their egocentric and socio-centric tendencies. They use the intellectual tools, concepts and principles, that critical thinking provides to enable them to analyze, assess and improve thinking, and then lead to better decisions. These people can work diligently to develop the intellectual virtues of intellectual integrity, intellectual civility, intellectual empathy, intellectual humility, intellectual sense of justice and confidence in reasoning and thinking.

In a seminal study on critical thinking and education in 1941, Edward Glaser defines critical thinking as "the ability to think critically", and involves three characteristics (1) an attitude of being disposed to consider in a thoughtful way the problems and subjects that come within the range of one's experience, (2) knowledge of the methods of logical inquiry and reasoning, and (3) some skills in applying those methods. Critical thinking calls for a persistent effort to examine any belief or supposed form of knowledge in the light of the evidence that supports it, and the further conclusions to which it tends.

Overall, critical thinking focuses on learning and developing the habitual intention to be truth-seeking, open-minded, systematic, analytical, confident in reasoning and prudent in making judgments. Failure to recognize the importance of critical thinking dispositions can lead to various forms of self-deception and closed-mindedness, both individually and collectively (Hindery, 2001).

Confirmation Bias / Myside Bias

In the critical thinking and decision making literature, the ability to deal with evidence in an unbiased way and the ability to accept multiple perspectives during problem solving and decision making are reasoning skills that are deemed preeminent important (Baron, 1991, 2000; Evans, 2002; 2007; Kuhn, 1999, 2001; Nickerson, 2004; Norris & Ennis, 1989; Perkins, 1995; Toplak & Stanovich, 2002; Wade & Tavris, 1993). Nevertheless, literature has amply shown that people are inclined to evaluate evidence, generate arguments and test hypotheses in a manner biased towards their own prior beliefs and opinions (Baron, 1991, 1995; Greenhoot, Semb, Colombo, & Schreiber, 2004; Kuhn, 1991; Perkins, 1985; Perkins, Faraday, & Bushey, 1991; Toplak & Stanovich, 2002; Voss, et al., 1991), and also people have difficulty accepting conclusions that conflict with what they think they know about the world (Evans, 2002; Evans, Barston, & Pollard, 1983; Evans, Newstead, Allen, & Pollard, 1994; Klauer, Musch, & Naumer, 2000). Nickerson (1998) reviews the related literature and terms these effects "confirmation bias", referring to "inappropriate bolstering of hypotheses or beliefs whose truth is in question" (p.175). He also attributes this inappropriate bolstering to prior

belief (the same as in "myside bias") or prior knowledge (the same as in "belief bias effect" in syllogistic reasoning, see Klauer et al 2000). Confirmation bias is perhaps the best known and most widely accepted notion of inferential error in the human reasoning literature (Evans 1989, p.41). In this series of studies, I use "myside bias" to refer to the prior belief effect in confirmation bias.

Confirmation bias connotes a less explicit, less consciously one-sided case-building process to justify an already drawn conclusion. The fundamental assumption of this concept is that people can and do engage in case-building unwittingly, without intending to treat evidence in a biased way or even being aware of doing so (Nickerson, 1998). People's natural tendency seems to favor information or look for evidence that is directly supportive of hypotheses they favor. People show this bias when they gather or remember information selectively, or when they interpret the information in a biased way. This effect is stronger for controversial ethical issues. People usually prefer sources that affirm their existing attitudes and also tend to interpret ambiguous evidence as supporting their existing position.

Confirmation biases are effects in information processing, referring to any way in which people avoid rejecting a belief, whether in searching for evidence, interpreting it, or recalling it from memory (Nickerson, 1998). Studies have repeatedly found that people tend to test hypotheses in a one-sided way, by searching for evidence consistent with the hypothesis they hold at that time. Rather than searching through all the relevant evidence, people search for information that supports their hypothesis (Baron, 2000). Further, confirmation biases are not limited to the collection of evidence. Even if two individuals have the same information, the way they interpret this information can be

biased towards their belief. Finally, even if someone has sought and interpreted evidence in a neutral manner, they may still remember it selectively to reinforce their expectations. This effect is called "selective recall", "access-biased memory" or "confirmatory memory" (Hastie & Park, 2005).

Biased search, interpretation and memory contribute to overconfidence, attitude polarization in the face of the contrary evidence, belief perseverance, the irrational primacy effect and illusory correlation. Explanations for the observed biases include wishful thinking and the limited human capacity of information processing. Another explanation is that people are weighing up the costs of being wrong, rather than investigating in a neutral, scientific way. People seldom seek evidence naturally that would show a hypothesis to be wrong, because they understand this to be an effective way to show it to be right if, they believe, it really is right. Nickerson (1998) mentioned that the question of the extent to which the confirmation bias can be modified by training deserves more research.

Reasoning and the Heuristics and Biases Debate

Normative models embody standards for cognitive activity—standards that, if met, serve to optimize the accuracy of beliefs and the efficacy of actions. These models evaluate thinking and decision making in terms of thinkers' personal goals (Baron, 2000). For example, subjective expected utility theory serves as a normative model for the choice of actions. However, a substantial research literature accumulated over three decades indicate that human behavior often deviated from the performance considered

normative on many reasoning tasks (Baron, 1994; Evans, 1989; Evans, & Over, 1996; Kahneman, Slovic, & Tversky, 1982; Kahneman & Tversky, 1996; Osherson, 1995; Plous, 1993; Shafir, & Tversky 1995).

In contrast, descriptive models—accurate specifications of the response patterns of human beings and theoretical accounts of those observed response patterns in terms of psychological mechanisms—are the goal of most work in empirical psychology. Descriptive models are theories about how people normally think—for example, how we solve problems in logic or how we make decisions (Baron, 2000). Therefore, many descriptive models are expressed in heuristics, or rules of thumb for using in certain situations by human beings.

The modern concept of heuristic methods was devised by George Polya. Polya (1945) defines "heuristic reasoning" as "reasoning not regarded as final and strict, but as provisional and plausible only, whose purpose is to discover the solution of the present problem." Beginning with Kahneman and Tversky (1972), researchers have used the idea of heuristics to explain the departure from normative models. People develop heuristics due to their usefulness in some situations; however, the use of heuristics often leads to biases. In the 1970s and early 1980s, demonstrating the divergence of human behavior from normative models was a main theme of the so-called heuristics and biases literature (Kahneman, Slovic, & Tversky, 1982).

In the reasoning and decision-making literature, the interpretation of the gap between the descriptive and the normative models has been the contentious debate for four decades (Cohen, 1981; Evans, & Over, 1996; Gigerenzer, 1996; Kahneman, &

Tversky, 1996; Stanovich, 2000; Stein, 1996). The debate has arisen since some investigators tried to attribute this gap to "systematic irrationalities" of human being. To dispute this contention, numerous investigators have argued that there are other explanations for this discrepancy (Cohen, 1981; Stein, 1996). First, "performance errors"—a person fails to apply a rule, algorithm, or strategy due to a momentary and fairly random lapse in ancillary processes necessary to execute the strategy (e.g., temporary lack of attention, memory deactivation, and other sporadic information-processing mishaps, like distraction).

Second, "computational limitations"— acknowledging cognitive resource limitations leads to the idea of replacing normative models with prescriptive models (Baron, 1985; Bell, Raiffa and Tversky, 1988; Cherniak, 1986; Harman, 1995; Oaksford, & Chater, 1993, 1995; Simon, 1956, 1957; Stich, 1990). Prescriptive models specify how processes of belief formation and decision making should be carried out to maximize goal satisfaction, given the limitations of the human cognitive apparatus and the situational constraints (e.g., time pressure).

Third, "wrong fit of normative model and task"— in interpreting performance, researchers might be applying the wrong normative model to the task (Cosmides, & Tooby, 1996; Levi, 1983; Lopes, 1981; Macdonald, 1986), or the participant might have construed the problem differently and provide the normatively appropriate answer to a different problem setting (Adler, 1984; Hilton, 1995; Levinson, 1995; Margolis, 1987; Schwarz, 1996). Fourth, "individual differences"—although the average person in the reasoning tasks displayed the divergence from the normative models, however, some people did give the standard normative responses. Therefore, individual differences and

their patterns of covariance have implications for explanations of why some human behavior departs from normative models while others don't. The last explanation for the gap between the normative and descriptive models will be the focus of this study.

Individual Differences in Rational Thought and Biases

Researchers have tried to examine the individual difference variables in skills at decoupling and avoiding confirmation biases (Stanovich, & West 2007). The most discussed individual difference factors in this steam of studies are cognitive ability and thinking dispositions. Recently, researchers also include cognitive reflection test in their studies.

Cognitive Ability

Levels of analysis in cognitive theory have been discussed by numerous investigators (Anderson, 1990, 1991; Dennett, 1978, 1987; Horgan, & Tienson, 1993; Levelt, 1995; Newell, 1982, 1990; Oaksford, & Chater, 1995; Pylyshyn, 1984; Sterelny, 1990; Stanovich, 2000). In this study, I adopt the terminology used by Stanovich (2000). The first level of analysis is biological level that is inaccessible to cognitive theorizing. The second level of analysis is algorithmic level, concerned with the computational processes necessary to carry out a task. The third level of analysis is intentional level, dealt with the goals of the system, beliefs relevant to the goals, and the choice of action

that is rational given the system's goals and beliefs (Anderson, 1990; Bratman, Israel, & Pollack, 1991; Dennett, 1987; Newell, 1982, 1990; Pollock, 1995; Stanovich, 2000).

The cognitive psychologists work largely at the algorithmic level by showing that human performance can be explained by positing certain information-processing mechanisms in the brain (Stanovich, & West, 2008), and have focused on cognitive capacities that underlie psychometric intelligence, including perceptual speed, discrimination accuracy, working memory capacity, and the efficiency of the retrieval of information stored in short-term and long-term memory (Ackerman, Kyllonen and Richards, 1999; Carpenter, Just and Shell, 1990; Deary, 2001; Hunt, 1987, 1999; Kane, & Engle, 2002; Lohman, 2000; Sternberg, 1977, 1985, 2000; Unsworth, & Engle, 2005).

Ever since Spearman (1904) first discovered positive manifold, intelligence indicators have correlated with a plethora of cognitive/personality traits and thinking abilities that are almost too many to enumerate (e.g., Ackerman, Kyllonen and Richards, 1999; Deary, 2001; Deary, Whiteman, Starr, Whalley, & Fox, 2004; Lubinski, 2000, 2004; Lubinski, & Humphreys, 1997). Generally speaking, people with higher cognitive capacity have larger working memories, faster reaction times and are more susceptible to visual illusions (Jensen, 1998). The g factor refers to the component variance that is common to all tests of ability (Brody, 1992, 1997; Carroll, 1993, 1995, 1997, 2003; Jensen, 1998; Mackintosh, 1998), and is not a measure of specific knowledge, skills or strategies for problem solving. Instead, the g factor may reflect individual differences in information processing, that is, the capacity and efficiency of the mental processes by which knowledge and skills are acquired and used (Jensen, 1998). Therefore, even

researchers use different ability tests in various studies, they all measure the same thing—general intelligence (the g factor).

Some studies have shown that there is modest relationship between cognitive ability and the performance on a variety of tasks in the heuristics and biases literature (Bruine de Bruin, Parker, & Fischhoff, 2007; DeShon, Smith, Chan and Schmitt, 1998; Handley, Capon, Beveridge, Dennis, & Evans, 2004; Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Newstead, Handley, Harley, Wright, & Farrelly, 2004; Parker, & Fischhoff, 2005; Perkins, & Ritchhart, 2004; Sá, West, & Stanovich, 1999; Stanovich, & West, 1997, 1998, 1999, 2000; Toplak, & Stanovich, 2002).

In the series of Stanovich and West studies, they have operationalized cognitive capacity in terms of cognitive ability and aptitude tasks such as the Scholastic Aptitude Test (SAT), Raven Matrices, and various vocabulary and reading comprehension tests (Stanovich, 2000, 2003, 2009; Stanovich, & West, 1997, 1999, 2000, 2007, 2008; Toplak, West, & Stanovich, 2011). All are known to have high loadings on the g factor (Carpenter, Just, & Shell, 1990; Carroll, 1993, 1997; Lubinski, & Humphreys, 1997; Matarazzo, 1972) and have been linked to information processing indicators of efficient cognitive computation (Caryl, 1994; Deary, 1995; Deary, & Stough, 1996; Detterman, 1994; Dougherty, & Haith, 1997; Fry, & Hale, 1996; Hunt, 1987; Stankov, & Dunn, 1993; Vernon, 1990, 1993). They found that the moderate correlation between performance on the reasoning tasks and cognitive ability provides some explanation for the gap between the normative and descriptive models. People with higher cognitive ability perform better in reasoning tasks and have less thinking biases.

Thinking Dispositions

"When is good thinking?" draws attention to an important dimension of thinking—good timing—attempting the right kind of thinking at the right moment, which relates to how thinking gets activated or mobilized when needed (Perkins, & Ritchhart, 2004). Such an account lies at the heart of a dispositional view of thinking, looking not only to what kinds of thinking people are able to do well, but what kinds of thinking they are disposed to take (Baron, 1985; Dewey, 1922; Ennis, 1987; Facione, Sanchez, Facione, & Gainen, 1995; Perkins, Jay, & Tishman, 1993; Perkins, & Ritchhart, 2004; Ritchhart, 2002; Stanovich, 1999).

The term thinking dispositions has its roots in philosophy. Philosophy has traditionally defined a disposition as a capacity, tendency, potentiality, or power to act in a certain way (Honderich, 1995). In psychology, dispositions can be viewed as loosely analogous to such dispositional properties. A good thinker is disposed to look at both sides of the case while encounter a broad generalization and look for hidden assumptions when a problem was initially framed (Perkins, & Ritchhart, 2004).

In psychology, cognitive ability tests are widely taken to be the token of good thinking. Critics of these instruments often point out that these tests fail to assess many domains of psychological functioning that are imperative in thinking itself. The arguments made by numerous cognitive theorists depend strongly on differentiating the algorithmic from the intentional level of analysis (Anderson, 1990, 1991; Baron, 1985, 1994; Matthews, Davies, Westerman, & Stammeres, 2000; Dennett, 1978, 1987; Horgan, & Tienson, 1993; Levelt, 1995; Marr, 1982; Newell, 1982, 1990; Oaksford, & Chater,

1995; Pylyshyn, 1984; Sloman, 1993; Sterelny, 1990). The algorithmic level is concerned with the computational processes and information-processing operations necessary to carry out a task. In contrast, the intentional level of analysis deals with the goals of the system, beliefs relevant to the goals, and the choice of strategy that is rational given the goals and beliefs.

Drawing on information-processing models of cognition, Stanovich and West (1997) argued that cognitive capacities and thinking dispositions "map onto different levels of analysis in cognitive theory. Variation in cognitive ability refers to individual difference in the efficiency of processing at the algorithmic level. In contrast, thinking disposition indexes individual differences at the intentional level" (p.9).

In individual differences studies, researchers have long recognized the distinction between cognitive capacities (general intelligence) and thinking dispositions/cognitive styles (Baron, 1985, Cacioppo, Petty, Feinstein, & Jarvis, 1996; Dole, & Sinatra, 1998; Ennis, 1987; Klaczynski, Gordon, & Fauth, 1997; Klaczynski & Lavallee, 2005; Kruglanski & Webster, 1996; Moshman, 1994; Nickerson, 2004; Norris, 1992; Parker & Fischhoff, 2005; Perkins, 1995; Schommer-Aikins, 2004; Sinatra & Pintrich, 2003; Stanovich, 1999, 2002, 2003; Stanovich & West, 1997, 1998, 1999, 2007, 2008; Sternberg, 1997, 2003; Toplak, West, & Stanovich, 2011). Thinking dispositions are concerned with beliefs, belief structure and attitudes towards forming and changing beliefs.

Thinking dispositions (Baron, 1988; Ennis, 1987; Perkins, 1995; Stanovich, & West, 1997, 1998, 1999, 2007, 2008; Toplak, West, & Stanovich, 2011) is used in our

study, although other researchers prefer terms such skill as intellectual style (Sternberg, 1988, 1989), cognitive emotions (Scheffler, 1991), habits of mind (Keating, 1990), inferential propensities (Kitcher, 1993), epistemic motivations (Kruglanski, 1990), constructive metareasoning (Moshman, 1994), and cognitive styles (Messick, 1984, 1994). They all refer to relatively stable psychological mechanism and strategies that tend to generate characteristic behavioral tendencies and tactics (Buss, 1991). Thinking dispositions, compared to cognitive capacities, are more malleable and teachable (Baron, 1985, 1988, 1993; Stanovich, 1999).

Baron (1985, p.15) proposed that "although you cannot improve working memory by instruction, you can tell someone to spend more time on problems before she gives up, and if she is so inclined, she can do what you say." He also defined rational thinking dispositions as things like "the disposition to weigh new evidence against a favored belief heavily (or lightly), the disposition to spend a great deal of time (or very little) on a problem before giving up, or the disposition to weigh heavily the opinions of others in forming one's own".

Actively Open-minded Thinking

Some studies have revealed that thinking dispositions associated with epistemic regulation (e.g., actively open-minded thinking) are related to individual differences in myside bias (Klaczynski, Gordon, & Fauth, 1997; Klaczynski, & Lavallee, 2005; Klaczynski, & Robinson, 2000; Sá, Kelley, Ho, & Stanovich, 2005). In a series of Stanovich and West studies (1997, 1998, 1999), they use a thinking disposition

composite score, which comprises several scales from their previous research or other research (e.g., openness, flexible thinking, counterfactual thinking, dogmatism, absolutism, categorical thinking, belief identification, etc.). Because several of the scales displayed moderate inter-correlations, a composite actively open-minded thinking score was formed by summing the scores on the Flexible Thinking, Openness-Ideas, and Openness-Values and subtracting the sum of the Absolutism, Dogmatism, and Categorical Thinking scales. They found that the composite score of thinking dispositions was associated with superior performance on rational thinking tasks, even after partialling out the variance accounted for by several measures of general cognitive ability. Thus, high scores on the AOT composite indicate openness to belief change and cognitive flexibility, while low scores on the AOT composite indicate cognitive rigidity and resistance to belief change.

Need For Cognition

Another frequently used scale of thinking disposition is Need for Cognition Scale (NCS), defined by the continuum of individual differences that stretches from chronic cognitive misers to cognitive cognizers. It refers to an individual's tendency to engage in and enjoy effortful cognitive endeavors (Cacioppo, Petty, Feinstein, & Jarvis, 1996). Contemporary research on NCS began with Cacioppo and Petty's (1982) proposal that there are stable dispositions in people's tendency to engage in effortful cognitive activity when given a task or making sense of the world and in their tendency to enjoy (or be less stressed by) cognitively effortful problems, life circumstances, or tasks. Individuals high

in need for cognition are naturally tend to seek, acquire, think about, and reflect back on information to make sense of stimuli, relationships, and events. On the contrary, individuals low in need for cognition are more likely to rely on others (e.g., celebrities and experts), cognitive heuristics, or social comparison processes (Cacioppo, & Petty, 1984; Cacioppo, Petty, & Morris, 1983).

Dual Process Theories

There are two opposite streams of research supporting opposite views in individual differences in debiasing ability. Stanovich and his colleagues (Stanovich, & West, 1997, 1998; Kokis, Macpherson, Toplak, West, Stanovich, 2002; Sá, West, & Stanovich, 1999) found moderate correlations between cognitive ability and the ability to avoid myside biases. On the other hand, Klaczynski and his colleagues (Klaczynski, 1997; Klaczynski, & Gordon, 1996; Klaczynski, Gordon, & Fauth, 1997; Klaczynski, & Lavallee, 2005; Klaczynski, & Narasimham, 1998; Klaczynski, & Robinson, 2000) have found that verbal ability (as measure of cognitive ability) was not correlated with the magnitude of myside bias effect. However, verbal ability was related to the quality of reasoning in both opinion-consistent and -inconsistent conditions.

Stanovich and West (2007) tried to figure out what made this discrepancy among these two streams of research, and proposed that the critical difference in experimental procedure may provide an explanation for it. The key factor is whether the researchers give the instructions to the participants to decouple from their prior knowledge and belief in the tasks. Stanovich and his colleagues strongly emphasized the necessity of

decoupling prior knowledge and belief from the tasks in the experiments. However, the Klaczynski Group didn't provide this instruction to their participants.

Moreover, many investigators have noted that within-subjects manipulations may draw the participants' attention to the variable of interest since this method asks what subjects think they should do and whether they think that outcome ought to matter (Fischhoff, Slovic, & Lichtenstein, 1979; Kahneman, 2000; Kahneman, & Frederick, 2002; Kahneman, & Tversky, 1996; LeBoeuf, & Shafir, 2003). Therefore, Stanovich and West (2007) argued that explicit instructions to decouple and within-subjects experiments sensitize the participant that detaching from prior opinions is a necessary requirement of the task. However, people of differing cognitive abilities have different levels of computational power available for the override operations that make decoupling possible (Stanovich, 1999, 2004; Stanovich, & West, 2000). Such an explanation is consistent with theoretical views emphasizing the difference between typical information processing and maximal information processing (Ackerman, 1994, 1996; Ackerman, & Heggestad, 1997; Ackerman, & Kanfer, 2004; Cronbach, 1949; Matthews, Zeidner and Roberts, 2002). Optimal performance are the situations that people are instructed to maximize their performance, while typical performance are unconstrained situations that people try their best without any overt instructions.

Stanovich and West(2007) proposed that in natural myside bias situations (without the instructions), bias is largely a function of the memeplexes subject to call from System 1 and is little moderated by decontextualizing operations by System 2. In contrast, when people are instructed to decouple from prior beliefs, their proclivities to do so (intentional level, e.g., thinking dispositions, like actively open-minded thinking and

need for cognition) and their capacity to do so (algorithmic level, e.g., general intelligence, like cognitive ability) come to the fore and predict the degree of myside bias.

Stanovich (1999) called this dual processes as "System 1"—largely automatic, unconscious and relatively undemanding of computational capacity, and conjoining properties of automaticity and heuristic processing, and "System 2"—largely conscious and demanding of computational capacity, and conjoining various characteristics of controlled processing, which allows us to process abstract concepts, to deliberate, to plan ahead, and to consider options carefully.

Cognitive Reflection Test

A framework, provided by the cognitive reflection test, of an incorrectly primed initial response that must be overridden fits in nicely with dual-process models (De Neys, & Glumicic, 2008; Evans, 1984, 2008, 2010; Evans, & Frankish, 2009; Kahneman, & Frederick, 2002; Lieberman, 2007, 2009; Sloman, 1996, 2002; Stanovich, 1999, 2009, 2011; Toplak, West, & Stanovich, 2011). The CRT carries properties across the boundary of the distinction between cognitive abilities and thinking dispositions. This distinction follows from separating optimal/maximal performance and typical performance.

The Cognitive Reflection Test (CRT) is a simple, three-item measure introduced by Frederick (2005). The three items are "easy" in the sense that their solution is easily understood when explained, yet reaching the correct answer often requires the suppression of an erroneous answer that pops up to mind right after reading the question.

Specifically, this test is designed to measure the tendency to override an incorrect prepotent response, and to engage in further reflection that leads to the correct response.

The quintessential item from the CRT was first proposed by Kahneman and Frederick (2002). They reframed the heuristics-and-biases literature in the concept of attribute substitution, a common mechanism used to lighten cognitive load. Attribute substitution occurs when a person actually needs to assess attribute A, but turns out to find that assessing attribute B (which is correlated with A) is cognitively easier, and finally he uses B instead of A. One of the problems they proposed in that paper is "A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?" People usually give the first response (10 cents) that comes to mind without thinking further, and the answer would not be right. When people answer this problem, many of them behave like cognitive misers, which mean they could not resist to the intuitive answer that pops up first (Dawes, 1976; Simon, 1955, 1956; Stanovich, 2009; Taylor, 1981; Tversky, & Kahneman, 1974). Frederick (2005) found that cognitive ability is no guarantee against this error (Toplak, West, & Stanovich, 2011).

Humans are cognitive misers due to their instinctive tendency to use heuristic processing mechanisms causing lower computational expense. Heuristic processes often provide a quick solution; however, modern life often requires deeper and thorough thinking. Therefore, being cognitive misers will sometimes impede people from being successful. Many effects in the heuristics-and-biases literature are the results of the human tendency to prefer miserly processing. Therefore, the CRT test could be a good predictor since an incorrect answer is initially primed and participants need to override

the miserly processing and replace it by a superior response (Toplak, West, & Stanovich, 2011).

The CRT taps both a cognitive ability and a thinking disposition dimension. On the one side, the CRT has moderate overlap with measures of cognitive ability (Frederick, 2005; Obrecht, Chapman, & Gelman, 2009; Toplak, West, & Stanovich, 2011). On the other side, the CRT also implicates thinking dispositions, particularly those related to reflectivity, the tendency to seek alternative solutions (Toplak, West, & Stanovich, 2011).

Logical Reasoning Ability

The biggest concern for the cognitive reflection test (CRT) is that the test only includes three questions and those questions require math ability. Therefore, I try to solve this issue and try to find the other type of questions that could have similar function as cognitive reflection test and do not require math ability. Also, these sets of questions need to have been proven showing consistent effects in both cognitive ability and thinking dispositions in a large sample size.

Logical reasoning skills are a complex weave of abilities that help you understand other people's sayings, evaluate the statements, decide what statement to accept or reject, explain a complicated idea, apply conscious quality control as you think, decide what or what not to do, etc. Logical reasoning ability is not a biological natural ability. Basically, it is a skill that can be learned and improved. Therefore, logical reasoning skill is the potentially perfect measurement for the current studies.

In the individual difference literature, few studies used tests or scales to measure logical reasoning ability. Therefore, I took one set of questions from logical reasoning sections in Law School Admission Test (LSAT) as questions of logical reasoning ability.

Arguments are a fundamental part of the law, and analyzing arguments is a key element of legal analysis. Training in the law builds on a foundation of basic reasoning skills. Law students must draw on the skills of analyzing, evaluating, constructing, and refuting arguments. They need to be able to identify what information is relevant to an issue or argument and what impact further evidence might have. Also, they need to be able to reconcile opposing positions and use arguments to persuade others.

Logical Reasoning questions evaluate the ability to analyze, critically evaluate, and complete arguments as they occur in ordinary language. The questions are based on short arguments drawn from a wide variety of sources, including newspapers, general interest magazines, scholarly publications, advertisements, and informal discourse. Therefore, these questions do not require math ability like CRT does.

The LSAT contains two logical reasoning ("LR") sections, commonly known as "arguments", designed to test the taker's ability to dissect and analyze arguments. Each question begins with a short argument or set of facts. This is followed by a prompt asking the test taker to find the argument's assumption, to select an alternate conclusion to the argument, to identify errors or logical omissions in the argument, to find another argument with parallel reasoning, or to choose a statement that would weaken or strengthen the argument. The questions are designed to assess a wide range of skills involved in thinking critically. Therefore, these questions are perfect for the current

studies to evaluate people's both thinking ability and thinking dispositions. I called this individual difference as "Logical Reasoning Ability" (LRA).

CHAPTER 3

STUDY ONE

Introduction

In an uncertain world, people love to feel being right since it helps us make sense of things. Psychologists also think it's akin to a basic drive. People frequently search for information that confirms their initial view of the issue and ignore the evidence that does not support their view. We do it automatically, usually without realizing. We do it partially because it seems easier to fit the new pieces into the picture-puzzle we are working on now, rather than imagining another new picture. It also helps shore up our visions of ourselves as accurate, right-thinking, consistent people who know things. Psychologists call it as confirmation bias.

Confirmation biases are effects in information processing, referring to any way in which people avoid rejecting a belief, whether in searching for evidence, interpreting it, or recalling it from memory (Nickerson, 1998). Studies have repeatedly found that people tend to test hypotheses in one-sided way, by searching for evidence consistent with the hypothesis they hold at that time (Baron, 2000). Even if someone has sought and interpreted evidence in a neutral manner, they may still remember it selectively to reinforce their expectations. This effect is called "selective recall", "access-biased memory" or "confirmatory memory" (Hastie & Park, 2005).

In the first study, I focus on the last component of confirmation bias- "confirmatory memory", which means people who suffer from confirmatory memory

would have trouble of recalling the opposite-side arguments and only generate the arguments on their side to support their own positions while being asked to list out all the reasons for them to form their positions. The tendency to give more attention and weight to the positive and the confirmatory has been shown to influence memory. When digging into our memories for data relevant to a position, we are more likely to recall data that confirms the position.

Nickerson (1998) said that the question of the extent to which the confirmation bias can be modified by training deserves more research. Also, Meliorists believe that everyday reasoning falls far short of the ideal, but can be improved through experience and education (Nisbett, 1993). In the past decade, Stanovich and West have provided interesting evidence on this point in support of the Meliorists. They observe that there is always a subset of decision makers giving a normative response on a decision task, indicating that normative cognitive strategies are not impossible to be followed by human beings, but are systematically held and used by some individuals. Researchers have tried to examine the individual difference variables in skills at decoupling and avoiding confirmation biases (Stanovich & West, 2000, 2007). The most discussed individual differences are cognitive ability and thinking dispositions. However, in the literature, most studies only used ANOVA to examine each factor using medium split one by one. In this study, I collected all the relevant individual difference factors and threw all factors in one multiple regression model. Then, I tried to find the best model and kept these factors for this study and future research.

"Consider the opposite" is a simple, straight forward, easy-to-execute debiasing strategy, and has received some attention in the literature. However, there is no

consistent and powerful evidence to show that debiasing strategy is useful. In this study, I prompt subjects to think harder and consider the opposite couple times to see if this individual debiasing strategy could work for the majority or could only work for some people. Therefore, I include all the most relevant individual difference variables in the model and try to find the best predictor of thinking ability and the ability of avoiding confirmation bias/myside bias.

Conceptual Framework

Opposing streams of research support different views of the role of individual differences in the generation of strong arguments or rationales. On one side, Stanovich and his colleagues (1997, 1998) found moderate correlations between cognitive ability and the ability to avoid certain biases in reasoning. On the other hand, Klaczynski and his colleagues (1997; 2000) as well as other researchers (cf., Barnett & Ceci, 2002) have found that biases are unrelated to cognitive ability. However, cognitive ability is only one of many possible individual differences that might theoretically affect decision quality and its antecedent, argument quality. Additional individual differences that might matter include thinking dispositions (i.e., habitual patterns of thought), logical reasoning ability, and a host of traits (e.g., need for cognition, risk taking, superstitiousness, need for closure, etc.). Whatever the magnitude of these individual differences may be, they can be compared in magnitude to known task-characteristics that either bias or debias subjects' reasoning. In the first study, I tried to collect many possible relevant individual differences and threw them all in the model to see which factors are not related to

confirmation bias and thinking ability should be removed from the model in this study and further studies. The framework in Figure 1 is the final model in this study, and Table 1 shows the possible framework for future studies.

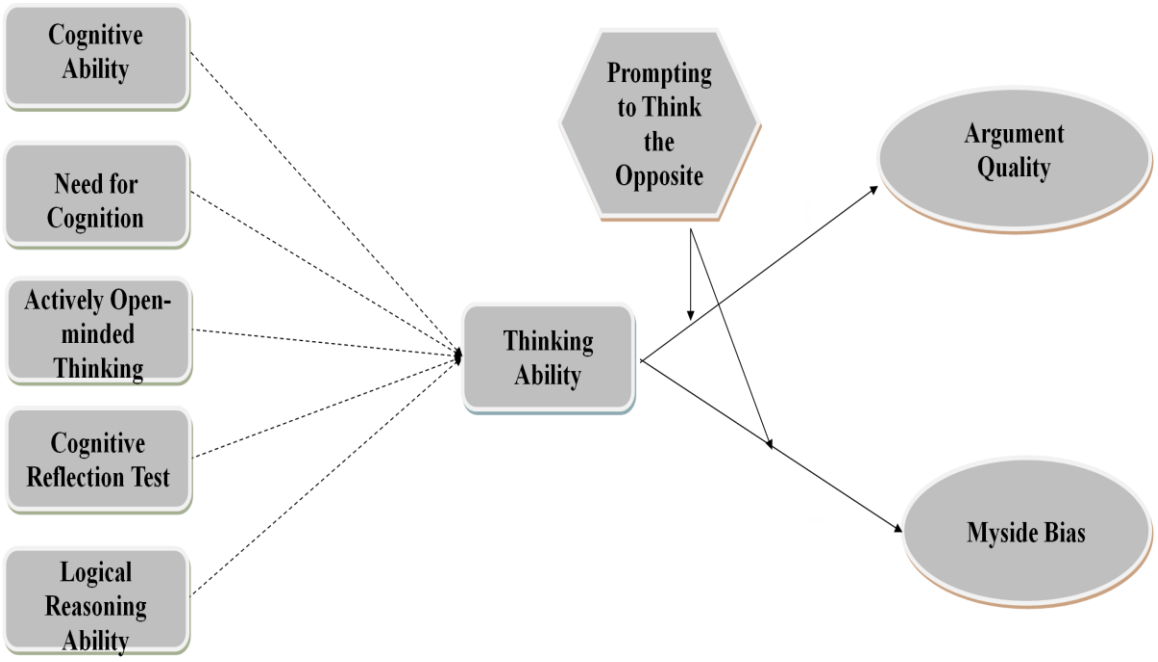

Figure 1. Study 1 Framework

Table 1. Possible Framework for Future Studies

Individual Characteristics	Cognitive Measures of Expertise	Potential Benefits
Cognitive Ability Active Open-minded Thinking Need for Cognition Cognitive Reflection Test Logical Reasoning Ability	Memory Knowledge Calibration Cognitive Effort Cognitive Structure Analytic Processing	Complete and Perfect Information Optimal Decision Making Lower Costs of Search, Product Usage & Consumption Higher Benefits from Search, Product Usage & Consumption

Cognitive Ability

Opposing streams of research support different views of the role of individual differences in the generation of strong arguments or rationales. On one side, Stanovich and his colleagues (1997, 1998) found moderate correlations between cognitive ability and the ability to avoid certain biases in reasoning. On the other hand, Klaczynski and his colleagues (1997; 2000) as well as other researchers (cf., Barnett & Ceci, 2002) have found that biases are unrelated to cognitive ability. However, Stanovich and West (2007) used the dual-process theories to explain the discrepancy between these two streams of research. They proposed that cognitive ability and thinking dispositions measures are negatively correlated with myside bias, when researchers provide respondents

instructions to decouple prior belief in the tasks. According to mixed findings in the literature, I hypothesize that,

H1a: Cognitive ability is not related to the argument generation quality.

H1b: Cognitive ability is not related to the myside bias.

H1c: The prompting of considering the other-side arguments moderates the correlation between cognitive ability and the argument generation quality.

H1d: The prompting of considering the other-side arguments moderates the correlation between cognitive ability and the myside bias.

Thinking Dispositions

Stanovich and his colleagues found that several thinking disposition measures (e.g., actively open-minded thinking, need for cognition scale, etc) are negatively related to myside bias (Stanovich, & West, 1997, 1998; Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Sá, West, & Stanovich, 1999). However, Stanovich and West later found that thinking dispositions are independent of natural myside bias (Stanovich, & West, 2007, 2008; Stanovich, 2009). They used dual-process theories to provide explanations for this discrepancy among these findings. They argued that some thinking disposition measures are still negatively correlated with myside bias after providing the

instructions of decoupling prior belief in the tasks. I have two measurements for thinking dispositions–actively open-minded thinking scale (AOT) and need for cognition scale (NCS). According to the literature, I hypothesize that,

H2a: AOT is not related to the argument generation quality.

H2b: AOT is not related to the myside bias.

H2c: The prompting of considering the other-side arguments moderates the correlation between AOT and the argument generation quality.

H2d: The prompting of considering the other-side arguments moderates the correlation between AOT and the myside bias.

H3a: NCS is not related to the argument generation quality.

H3b: NCS is not related to the myside bias.

H3c: The prompting of considering the other-side arguments moderates the correlation between NCS and the argument generation quality.

H3d: The prompting of considering the other-side arguments moderates the correlation between NCS and the myside bias.

Logical Reasoning Ability

The law school admission test (LSAT) is a univariate test designed to measure reasoning ability (Henderson, 2004). The logical reasoning section in LSAT is designed to measure skills, including the reading and comprehension of complex texts with accuracy and insight, the organization and management of information and the ability to draw reasonable inferences from it, the ability to think critically, and the analysis and evaluation of the reasoning and arguments of others. Hence, I propose that people who are high in logical reasoning ability (LRA) will perform better in providing good arguments for their constructed preferences and will be able to generate more other side arguments. Therefore, I hypothesize that,

H4a: LRA is positively related to the argument generation quality.

H4b: LRA is negatively related to the myside bias.

H4c: The prompting of considering the other-side arguments moderates the correlation between LRA and the argument generation quality.

H4d: The prompting of considering the other-side arguments moderates the correlation between LRA and the myside bias.

Cognitive Reflection Test

The CRT would seem to be ideally constructed as a good predictor of performance on heuristic-and-biases tasks; however, the results have been inconsistent. Frederick (2005) observed that CRT could predict performance on temporal discounting, the tendency to choose high-expected-value gambles and framing effects. Also, Cokely and Kelley (2009) found moderate correlation between CRT and the proportion of choices consistent with expected value. Oechssler, Roider, and Schmitz (2009) found that CRT was related to the number of expected-value choices and the tendency to commit the conjunction fallacy. Koehler and James (2010) found significant correlations between CRT and the use of maximizing strategies on probabilistic prediction tasks. Toplak, West and Stanovich (2011) showed that the CRT could explain substantial variance in rational thinking performance after partialling out cognitive ability, thinking dispositions, and executive functioning. Furthermore, they also demonstrated that the CRT actually is better predictor than intelligence, thinking dispositions and executive functioning.

On the contrary, Campiteli and Labollita (2010) found little relation between CRT and the choice of high-expected-value gambles. Obrecht, Chapman and Gelman(2009) found no relation between CRT and the degree of encounter frequency bias. According to the mixed results from the literature, I hypothesize that,

H5a: CRT is not related to the argument generation quality.

H5b: CRT is not related to the myside bias.

H5c: The prompting of considering the other-side arguments moderates the correlation between CRT and the argument generation quality.

H5d: The prompting of considering the other-side arguments moderates the correlation between CRT and the myside bias.

Methodology

Procedure

Subjects comprised 115 students in a major northeastern university in the U.S. who participated both online and in a lab setting for course credits. In all cases, data was collected using a computerized Qualtrics survey. Subjects were randomly presented with one of five consumption-related issue (e.g., “buying extended warrantees is a good idea.”, "going to college is a good idea", "for most products, higher price indicates higher quality (you 'get what you pay for')", "generally, online reviews of products and services are trustworthy", "products from large, name-brand companies are better than those from smaller companies"). Each issue had a number of possible pro and con arguments associated with it. First, subjects indicated their “spot” agreement with the statement and provided reasons for their view in an open-ended fashion. Prior research demonstrated that most subjects are prone to myside bias (arguing in a one-sided fashion for a chosen

view). Then, all subjects were provided with successively stronger prompts to consider all possible views about the argument, and to challenge their initial views.

After generating their argument, I collected a battery of individual difference measures comprising numerous categories (SAT Score, Actively Open-Minded Thinking, Need for Cognition, Risk Seeking, Belief in Superstitions, Conscientiousness, etc.), as well as objectively scored questions (e.g., Logical Reasoning Ability, Cognitive Reflection Test). In the end, bunch of demographics information were collected.

Measurements

Dependent Variables

Overall quality. Arguments that generated by respondents were independently graded for quality by three Ph.D. students in philosophy who routinely teach courses in critical thinking and informal reasoning, and who used a consensus rubric to guide their evaluations. The argument quality score ranges from 0 to 100, meaning better quality for higher scores. Then, I took the average of all three scores for each argument quality measurement.

Myside bias. My-side and other-side arguments were independently counted by three Ph.D. students in philosophy who routinely teach courses in critical thinking and informal reasoning, and who used a consensus rubric to guide their evaluations. Then, I took the average of all my-side and other-side counts by three philosophy Ph.D. students.

According to the definition of myside bias, the myside bias score equals to myside argument counts minus otherside argument counts (Toplak & Stanovich, 2003).

Independent Variables

Cognitive ability. Respondents were asked to provide their reading, mathematical, writing, total SAT scores, and how many times they took SAT in the demographics section. A indication of the validity of such self-reported scores was obtained by several studies (Stanovich & West, 1998; West & Stanovich, 1991). Also, my reliance on the total SAT score as a measure of cognitive ability is justified by the fact that it loads highly on psychometric g (Frey & Detterman, 2004), and that it is strongly associated with working memory—the quintessential indicator of computational capacity in cognitive science (Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Conway, Kane, and Engle, 2003; Engle, 2002; Engle, Tuholski, Kaughlin, & Conway, 1999; Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle, 2004). The mean reported reading SAT score of the respondents was 543 (SD_92), the mean reported mathematical SAT score was 549 (SD_93), the mean reported writing SAT score was 543 (SD_94), and the mean total SAT score was 1636 (SD_227). Differently from the most literature, I treated the total SAT score as a continuous variable instead of dichotomized variable, indicating higher cognitive ability for higher total SAT score.

Actively open-minded thinking scale. I adopted the actively open-minded thinking scale from Stanovich and West (2007). This scale was composed of 41 items drawn from a variety of sources as follows: 8 items from the Openness-Values facet of

the Revised NEO Personality Inventory (Costa & McCrae, 1992), 10 items from a flexible thinking scale developed by Stanovich and West (1997), 9 items from the belief identification scale developed by Sá, West, and Stanovich (1999), 9 items measuring dogmatism (Paulhus, & Reid, 1991; Robnison, Shaver, & Wrightsman, 1991; Troldahl, & Powell, 1965), 3 items from the categorical thinking subscale of Epstein and Meier's (1989) constructive thinking inventory, and 2 items from a counterfactual thinking scale developed by Stanovich and West (1997). All items were scored in the direction that higher scores represented a greater tendency toward open-minded thinking. Examples of items are "Certain beliefs are just too important to abandon no matter how good a case can be made against them" (reverse scored), "People should always take into consideration evidence that goes against their beliefs", "No one can talk me out of something I know is right" (reverse scored). The response format for each item in the scale was: Strongly Agree (7), Moderately Agree (6), Slightly Agree (5), Moderate (4), Slightly Disagree (3), Moderately Disagree (2), and Strongly Disagree (1). The score of AOT scale was obtained by summing the responses of the 41 items (Mean_185.6, SD_23). I treated the AOT score as a continuous variable, indicating higher actively open-minded thinking for higher AOT score.

Need for cognition scale. The need for cognition scale is an assessment instrument that quantitatively measures "the tendency for an individual to engage in and enjoy thinking" (Cacioppo & Petty, 1982). Cacioppo and Petty collaborated with Chuan Feng Kao to shorten the scale to the 18-item format, which is popularly used in studies nowadays (Bornstein, 2004; Gauthier, Christopher, Walter, Mourad, & Marek, 2006; Coutinho & Woolery, 2004; Sadowski & Gülgös, 1992, 1996). The 18-item need for

cognition scale (Cacioppo, Petty, & Kao, 1984) is adopted in this study. Sample questions are "Thinking is not my idea of fun" (reverse scored), "The notion of thinking abstractly is appealing to me", "I find satisfaction in deliberating hard and for long hours." The response format for each item in the scale was: Strongly Agree (7), Moderately Agree (6), Slightly Agree (5), Moderate (4), Slightly Disagree (3), Moderately Disagree (2), and Strongly Disagree (1). The score of NCS was obtained by summing the responses of the 18 items (Mean_81.7, SD_12.84). I treated the NCS score as a continuous variable, indicating higher thinking needs for higher NCS score.

Logical reasoning ability. I took a set of questions from logical reasoning section in LSAT and did a pretest to remove some extreme difficult and simple questions among 105 students. After pretest, I kept six questions for the study based on the results from the Item Response Analysis. Typically, the primary goal of IRT is to examine the quality of items. Also, one can use the ability estimates yielded from IRT modeling to differentiate people's ability. First, I graded the respondents' answers as correct (1) or wrong (0) and then put these scores of six questions all together into Item Analysis. Second, item analysis showed the item characteristic curves (ICC), indicating the probability of answering the item correctly at different levels of respondent ability, whereas the item information function (IIF), informing us how much reliable information about the respondent I can obtain at different levels of student ability. Third, IRT gave me the respondents' ability scores based on how many questions they answered correctly and also based on how difficult those questions were. Then, I used these ability scores as logical reasoning ability (LRA) scores (Mean_-0.0762, SD_0.9025).

Cognitive reflection test. The Cognitive Reflection Test (CRT) is a three-item measure introduced in the journal literature by Frederick (2005). This test is designed to measure the tendency to override a prepotent response that is incorrect and to engage in further reflection that leads to the correct response. Sample question is as follows: A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost? I graded the answers as correct (1) or wrong (0), and then added up the scores of these three questions. I treated CRT score as a continuous variable, ranging from 0 to 3 (Mean_0.6625, SD_0.8992).

Results

Argument Quality

In Table 2, Standard Least Squares Analysis showed that logical reasoning ability was statistically significant ($p < 0.05$) and indicated that the logical reasoning ability was the best predictor of performance on argument quality. The logical reasoning ability was the variable that predicted the most unique variance (11.35%). This analysis showed that logical reasoning ability is the best predictor of performance on argument quality before prompting to think about the other-side arguments. Therefore, H1a, H2a, H3a, H4a, and H5a were all supported. This means that people who are higher in logical reasoning ability can generate better arguments to support their preferences, even before prompting to think about the other-side arguments. They can support their positions with more

sound arguments. However, all the other individual difference variables are not related to initial argument generation quality.

Table 3 showed that logical reasoning ability overall was better predictor of all argument quality through three prompting. However, there was no interaction effect between prompting and logical reasoning ability, maybe since people who were high in logical reasoning ability performed better in the initial argument; therefore, there were no significant difference before and after prompting. Overall, H1c, H2c, H3c, H4c, H5c were not supported, which means prompting to think the other side arguments didn't help changing the main effect of individual difference variables on argument generation quality.

Even, the slope of logical reasoning ability on argument generation quality was getting larger after each prompting ($b = 6.27 \rightarrow b = 6.75 \rightarrow b = 7.00 \rightarrow b = 7.32$), which means the correlation between logical reasoning ability and argument generation quality was getting stronger after each prompt. It showed that prompting to think harder and the other-side argument did help improving people's argument generation quality, however, it only worked on people with higher logical reasoning ability.

Table 2. SLS Analysis Predicting Initial Argument Generation Quality

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	22.3773	20.3746	1.10	0.2755
SAT	0.0109	0.0092	1.19	0.2390
AOT	0.0228	0.0961	0.24	0.8133
NCS	-0.1226	0.1651	-0.74	0.4601
LRA	6.2698	2.2221	2.82	0.0060**
CRT	-0.2965	2.3192	-0.13	0.8986

Note: *: p<0.1; **: p<0.05

Table 3. Repeated Measures MANOVA Analysis Predicting Argument Generation Quality

Parameter	Group Main Effect		Main Effect of Prompt Within Group	
	F	*(df1, df2)*	*F*	*(df1, df2)*
Intercept	1.33	(1, 78)	0.34	(3, 76)
SAT	1.67	(1, 78)	0.76	(3, 76)
AOT	0.09	(1, 78)	2.67	(3, 76)
NCS	0.85	(1, 78)	0.96	(3, 76)
LRA	9.16**	(1, 78)	0.97	(3, 76)
CRT	0.01	(1, 78)	0.79	(3, 76)

Note: *: p<0.1; **: p<0.05

Myside Bias

In Table 4 and Table 6, Standard Least Squares Analysis showed that no parameter was statistically significant ($p < 0.05$), and indicated that people are inclined to generate more myside arguments over otherside arguments before prompting to consider the other-side arguments. Therefore, H1b, H2b, H3b and H5b were all supported, except H4b. The results showed that all the individual difference variables were not related to initial myside bias, which meant everyone suffered from myside bias before prompting to consider the other-side arguments.

Table 5 and Table 7 showed that there was interaction effect between prompting to consider the other-side arguments and logical reasoning ability, which means people having higher logical reasoning ability could suppress myside bias inclination and think critically by generating more other-side arguments relative to myside arguments. Therefore, H1d, H2d, H3d, H5d were not supported, whereas H4d was supported. After prompting to consider the other-side arguments, people with higher logical reasoning ability can generate more other-side arguments and lead to avoid myside bias than people with lower logical reasoning ability. Most interestingly, prompting to consider the other-side arguments flipped the correlation between logical reasoning ability and the myside bias ($b = 0.04 \rightarrow b = -0.85 \rightarrow b = -0.32 \rightarrow b = -0.34$). This showed that before prompting, people with higher logical reasoning ability had more myside bias. However, after prompting, people with higher logical reasoning ability had less myside bias. Prompting changed the direction of the relationship between logical reasoning ability and the myside bias.

Table 4. SLS Analysis Predicting Initial Myside Bias

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	3.3957	1.5413	2.20	0.0305**
SAT	-0.0005	0.0007	-0.66	0.5110
AOT	-0.0023	0.0073	-0.31	0.7551
NCS	-0.0115	0.0125	-0.92	0.3586
LRA	0.0425	0.1680	0.25	0.8010
CRT	-0.0861	0.1754	-0.49	0.6251

Note: *: p<0.1; **: p<0.05

Table 5. Repeated Measures MANOVA Analysis Predicting Myside Bias

Parameter	Group Main Effect		Main Effect of Prompt Within Group	
	F	*(df1, df2)*	*F*	*(df1, df2)*
Intercept	5.95**	(1, 78)	1.21	(3, 76)
SAT	0.58	(1, 78)	0.03	(3, 76)
AOT	0.29	(1, 78)	0.36	(3, 76)
NCS	1.92	(1, 78)	1.31	(3, 76)
LRA	0.97	(1, 78)	3.06**	(3, 76)
CRT	0.17	(1, 78)	1.84	(3, 76)

Note: *: p<0.1; **: p<0.05

Table 6. SLS Analysis Predicting Initial Otherside Argument Counts

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0.3925	0.7938	-0.49	0.6224
SAT	0.0006	0.0004	1.67	0.0991*
AOT	-0.0004	0.0037	-0.12	0.9084
NCS	-0.0013	0.0064	-0.21	0.8372
LRA	0.0882	0.0865	1.02	0.3113
CRT	0.0269	0.0904	0.30	0.7669

Note: *: p<0.1; **: p<0.05

Table 7. Repeated Measures MANOVA Analysis Predicting Otherside Argument Counts

Parameter	Group Main Effect		Main Effect of Prompt Within Group	
	F	*(df1, df2)*	*F*	*(df1, df2)*
Intercept	0.05	(1, 78)	0.80	(3, 76)
SAT	1.66	(1, 78)	0.31	(3, 76)
AOT	0.28	(1, 78)	0.37	(3, 76)
NCS	0.0007	(1, 78)	1.18	(3, 76)
LRA	8.19**	(1, 78)	4.07**	(3, 76)
CRT	0.0003	(1, 78)	1.08	(3, 76)

Note: *: p<0.1; **: p<0.05

Conclusion

Interestingly, measures of intelligence, often taken as the benchmark of good thinking, do not assess the avoidance of myside bias and one-sided bias—the foundational skills in the critical thinking literature and in the rational thinking literature aimed at educational applications (Baron, 2000; Kuhn, 1993, 2005; Nussbaum & Sinatra, 2003; Perkins, 1995; Sternberg, 1997, 2001, 2003).

Furthermore, I will argue here that there are theoretical reasons to believe that these classes of thinking skills—myside thinking and one-sided thinking—are likely to be unusually dissociated from measures of cognitive ability. Because intelligence is such a ubiquitous associate of cognitive performance, it is noteworthy to identify the types of cognitive tasks that tend to eliminate its predictive power (Stanovich & West, 2008).

Surprisingly, I found little or no relationship between measures of cognitive ability and bias or argument quality prior to prompting. However, there was a strong positive relationship between acumen in logical reasoning ability and both argument quality and reduction of common biases in reasoning such as myside bias.

In addition, I found substantial effects of some dispositional measures on the degree to which debiasing strategies helped subjects. Debiasing effects moderated the main effect of logical reasoning ability on myside bias. Before prompting to use debiasing strategy, people with higher logical reasoning ability had more myside bias; however, the debiasing strategy helped people with higher logical reasoning ability suppress myside bias inclination and generated more other-side arguments. Therefore, the slope of logical reasoning ability on myside bias changed from positive to negative,

which meant that people with higher logical reasoning ability had less myside bias after generating and considering the other-side arguments.

Cognitive ability and some thinking disposition measures actually are not correlated with argument generation quality and myside bias, even after prompting, in consumption-related issues, which are inconsistent with prior research in non-consumption-related issues. However, logical reasoning ability is a good predictor of argument generation quality and debiasing.

Discussion

Our results showed little or no relationship between "innate" characteristics (e.g., cognitive ability, other traits) and argument quality. However, I showed strong, positive, effects of thinking dispositions on both quality and bias reduction. This is very encouraging, because thinking dispositions are more subject to conscious change, training, and the development of expertise than "trait" types of individual differences. In addition, I identified one dispositional moderators of the efficacy of debiasing techniques, which is a novel finding.

Other Interesting Findings and Future Study

In Study 1, I also observed some interesting findings about attitude change. Based on Table 8 and Table 9, the results showed that logic reasoning ability interacted with attitude polarization and confidence change. Before prompting to execute the debiasing strategy, the correlation between logical reasoning ability and attitude/confidence strength was negative, meaning that people with lower logical reasoning ability had stronger opinion and higher confidence. However, after doing the debiasing strategy, the correlation of logical reasoning ability and attitude/confidence strength changed from negative to positive (Attitude Change: b = -0.02 → b = 0.18; Confidence Change: b = -1.56 → b = 0.73), showing that people who were higher in logic reasoning ability had stronger opinion (attitude polarization) and higher confidence after they generated more both myside and otherside arguments. I proposed two possible mechanisms by which couple times of argument generation requests might degrade attitude and confidence for people low in logical reasoning ability and increase attitude and confidence for people high in logical reasoning ability. The first possible explanation is that people lower in logical reasoning ability are more susceptible to the illusion of explanatory depth. People generally tend to believe they are more competent than they actually are, and then lead to overconfidence and stronger opinions. Especially, the illusion of explanatory depth arises when people overestimate their ability to explain mechanical and natural processes (Rozenblit & Keil, 2002).

The second explanation is from motivated reasoning studies. The studies of motivated reasoning show that it makes sense to look for arguments for our opinions before we find ourselves called upon to state them. People will be ready and feel

confident if the search for arguments is successful. However, if the search for arguments fails, then it might be better to adopt a weaker position, one that is easier to defend. In Study 1, the results show that people who are low in logical reasoning ability believe that they understand these concepts quite deeply and are surprised by the shallowness of their own explanations when prompted to generate the arguments to support and oppose their positions. In contrast, people who are high in logical reasoning ability build their confidence and stronger opinions by generating and weighing more pros and cons after the prompting.

In attitude polarization literature, there are mixed findings. Further, according to the Argumentative Theory, people are all skilled arguers. People will be good at reasoning once they are in argumentative settings, and will lead to good solutions and depolarization. However, based on the findings of Study 1, I propose that the moderating effect of individual difference can provide some explanations to this stream of literature. I believe that not everyone is good thinker, even after using the debiasing strategy (e.g., prompting to consider the opposite side) (Kuhn, 1992) and arguing with other people will not always end up with good outcomes and depolarization. I will investigate this issue further in Study 2 and 3.

Table 8. Repeated Measures MANOVA Analysis Predicting Strong Opinion

Parameter	Group Main Effect		Main Effect of Prompt Within Group	
	F	*(df1, df2)*	*F*	*(df1, df2)*
Intercept	2.3807	(1, 78)	2.0135	(1, 78)
SAT	0.1559	(1, 78)	1.8797	(1, 78)
AOT	0.1757	(1, 78)	7.3379**	(1, 78)
NCS	1.0724	(1, 78)	0.0029	(1, 78)
LRA	0.7754	(1, 78)	5.3861**	(1, 78)
CRT	0.0112	(1, 78)	0.1219	(1, 78)

Note: *: p<0.1; **: p<0.05

Table 9. Repeated Measures MANOVA Analysis Predicting Confidence Level

Parameter	Group Main Effect		Main Effect of Prompt Within Group	
	F	*(df1, df2)*	*F*	*(df1, df2)*
Intercept	5.4903**	(1, 78)	1.4939	(1, 78)
SAT	0.6396	(1, 78)	0.2567	(1, 78)
AOT	0.4666	(1, 78)	2.2583	(1, 78)
NCS	2.7419	(1, 78)	0.6112	(1, 78)
LRA	0.0681	(1, 78)	3.8436*	(1, 78)
CRT	0.1902	(1, 78)	0.8841	(1, 78)

Note: *: p<0.1; **: p<0.05

CHAPTER 4

STUDY TWO

Introduction

In daily life, consumers are bombarded with tons of information about various issues, disputes and controversies, and then update their beliefs by discriminating, analyzing and confirming/disconfirming information. Statistics theory would suggest that more information is always better since relevant information can only increase the accuracy of judgments. However, people may not be able to discriminate the relevant information from the irrelevant ones (Peterson & Pitz, 1988), or they distort the information to favor their stances (confirmation bias, Nickerson, 1998) or the other plausible explanation.

> "It is the peculiar and perpetual error of the human understanding to be more moved and excited by affirmatives than by negatives."
>
> Francis Bacon

Confirmation bias refers to a type of selective attention whereby one tends to notice and to look for what confirms one's beliefs, and to ignore, not look for, or undervalue the relevance of what contradicts one's beliefs. It is a phenomenon that decision makers have been demonstrated to actively seek out and assign more weights to

evidence that confirms their prior beliefs, and ignore or underweigh evidence that could disconfirm their prior beliefs.

As such, it can be thought as a form of selection bias in collecting evidence. For example, if you believe that during a full moon, there is an increase in admissions to the emergency room in the hospital where you work, you will take special notice of admissions during a full moon. However, you will be inattentive to the admissions during other nights of the month. A tendency to do this repeatedly unjustifiably strengthens your belief in the correlation between the full moon and accidents in the local area.

The tendency to give more attention and weight to data that support our beliefs than to data that are contrary to our beliefs is especially pernicious when our beliefs are little more than prejudices. If our beliefs are firmly established on solid evidence and valid confirmatory experiments, the tendency to give more attention and weight to data that fit with our beliefs should not lead us astray as a rule. If we become blinded to evidence truly refuting a favored hypothesis, we have crossed the line from reasonableness and closed-mindedness.

Numerous studies have demonstrated that people, in general, give more value to confirmatory information to supportive data. The most likely reason for this phenomenon is that it is easier to deal with cognitively (Gilovich, 1993). It is way easier to see how a piece of data supports a position than to see how it disconfirms against our positions. Considering a typical Experimental Social Psychology experiment, data are easily managed to count as success, while negative instances require intellectual effort to

even see them as negative or to consider them as significant and include in the decision set.

Individuals have to continually remind themselves of this tendency and actively seek out or value data contrary to their beliefs. Since this is unnatural, it appears that the ordinary person is doomed to bias. However, in study 1, I found that people with higher logical reasoning ability is able to conquer this bias in argument generation task. In this study, I look into the second part of confirmation bias-biased interpretation. Biased interpretation happens when two individuals interpret the same information differently and biased towards their own beliefs. I would like to investigate which individual differences are the best predictors of argument evaluation ability and the ability of avoiding interpretation bias.

The other interesting question is people who are good at argument generation can perform well in argument evaluation tasks and also be able to avoid confirmation/disconfirmation bias while evaluating arguments? I found that people with higher logical reasoning ability can generate more both myside and otherside arguments, and then lead to less confirmation bias/myside bais in Study 1. In this study, I would like to investigate if the individual difference can be carried over into different domains and different tasks.

Conceptual Framework

Opposing streams of research support different views of the role of individual differences in the evaluation of strong arguments or rationales. On the one side, Stanovich and his colleagues (1997, 1998) found moderate correlations between cognitive ability and the ability to avoid certain biases in reasoning. On the other side, Klaczynski and his colleagues (1997; 2000) as well as other researchers (cf., Barnett & Ceci, 2002) have found that biases are unrelated to cognitive ability. However, cognitive ability is only one of many possible individual differences that might theoretically affect decision quality and its antecedent, argument quality. Additional individual differences that might matter include thinking dispositions (i.e., habitual patterns of thought), logical reasoning, cognitive reflection test (CRT) and a host of traits (e.g., need for cognition). Whatever the magnitude of these individual differences may be, they can be compared in magnitude to known task-characteristics that either bias or debias subjects' reasoning.

In the critical thinking literature, the ability to deal with evidence in an unbiased manner and the ability to take multiple perspectives when thinking about a problem are reasoning skills that are deemed pre-eminent importance to good thinking (Baron 1991, 2000; Evans, 2002, 2007; Kuhn, 1999, 2001; Nickerson, 2004; Norris & Ennis, 1989; Perkins, 1995; Toplak & Stanovich, 2002). However, research has demonstrated that people often fail to display critical thinking skills by avoiding confirmation bias: evaluating evidence, generating evidence and testing hypotheses in a manner biased towards their own opinions (Baron, 1991; Klaczynski & Robinson, 2000; Kuhn, 1991, 2005; Nickerson, 1998; Sá, Kelley, Ho, & Stanovich, 2005; Toplak & Stanovich, 2003) and one-sided thinking biases: preferring arguments that are one-sided rather than

balanced arguments that reflecting many different perspectives (Baron, 1991, 1995; Kuhn, 1991, 2001; Kuhn & Weinstock, 2002). I have gone through the literature review in individual differences in thinking biases in Study 1. Therefore, based on the literature review in introduction and first study, I have similar hypotheses for evaluation tasks in this study too.

Cognitive Ability

Opposing streams of research support different views of the role of individual differences in the evaluation of strong arguments or rationales. On one side, Stanovich and his colleagues (1997, 1998) found moderate correlations between cognitive ability and the ability to avoid certain biases in reasoning. On the other hand, Klaczynski and his colleagues (1997; 2000) as well as other researchers (cf., Barnett & Ceci, 2002) have found that biases are unrelated to cognitive ability. However, Stanovich and West (2007) used the dual-process theories to explain the discrepancy between these two streams of research. They proposed that cognitive ability and thinking dispositions measures are negatively correlated with confirmation bias, when researchers provide respondents instructions to decouple prior belief in the tasks. I do not give instructions of decoupling prior beliefs in the evaluation tasks. Also, in the first study, I found no relationship between quality of argument generation/confirmation bias and cognitive ability. According to the literature and the results from the first study, I hypothesize that,

H1.1: (a) Cognitive ability is not related to the argument generation quality; (b) the prompting of debiasing strategy does not moderate the main effect of cognitive ability on the argument generation quality.

H1.2: Cognitive ability is not related to the argument evaluation ability.

H1.3: (a) Cognitive ability is not related to the myside bias in argument generation task; (b) the prompting of debiasing strategy does not moderate the main effect of cognitive ability on the myside bias in argument generation task.

H1.4: Cognitive ability is not related to the myside bias in argument evaluation task.

Thinking Dispositions

Stanovich and his colleagues found that several thinking disposition measures (e.g., actively open-minded thinking, need for cognition scale, etc) are negatively related to myside bias (Stanovich, & West, 1997, 1998; Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Sá, West, & Stanovich, 1999). However, Stanovich and West later found that thinking dispositions are independent of natural myside bias (Stanovich, & West, 2007, 2008; Stanovich, 2009). They used dual-process theories to provide explanations for this discrepancy among these findings. They argued that some thinking disposition measures are still negatively correlated with myside bias after providing the

instructions of decoupling prior belief in the tasks, but not correlated with natural myside bias. Also, in the first study, I found no relationship between quality of argument generation/myside bias and thinking dispositions. According to the literature and the results from first study, I hypothesize that,

H2.1: (a) AOT is not related to the argument generation quality; (b) the prompting of debiasing strategy does not moderate the main effect of AOT on the argument generation quality.

H2.2: AOT is not related to the quality of argument evaluation task.

H2.3: (a) AOT is not related to the myside bias in argument generation task; (b) the prompting of debiasing strategy does not moderate the main effect of AOT on the myside bias in argument generation task.

H2.4: AOT is not related to the myside bias in argument evaluation task.

H3.1: (a) NCS is not related to the argument generation quality; (b) the prompting of debiasing strategy does not moderate the main effect of NCS on the argument generation quality.

H3.2: NCS is not related to the argument evaluation ability.

H3.3: (a) NCS is not related to the myside bias in argument generation task; (b) the prompting of debiasing strategy does not moderate the main effect of NCS on the myside bias in argument generation task.

H3.4: NCS is not related to the myside bias in argument evaluation task.

Logical Reasoning Ability

The law school admission test (LSAT) is a univariate test designed to measure reasoning ability (Henderson, 2004). The logical reasoning section in LSAT is designed to measure skills, including the reading and comprehension of complex texts with accuracy and insight, the organization and management of information and the ability to draw reasonable inferences from it, the ability to think critically, and the analysis and evaluation of the reasoning and arguments of others. Hence, I propose that people who are high in logical reasoning ability will perform better in evaluating arguments and will be able to differentiate the good ones from the bad ones, no matter which side they are supporting at. Although no studies support my proposition yet, logical reasoning ability is one of the potential cognitive ability and thinking disposition measures. However, in the first study, I found strong positive relationship between the initial quality of argument generation and logical reasoning ability, but the prompting of debiasing strategy did not moderate the main effect. Also, there was strong positive relationship between the decrease in myside bias and the logical reasoning ability after prompting. Therefore, I hypothesize that,

H4.1: (a) LRA is positively related to the argument generation quality; (b) the prompting of debiasing strategy does not moderate the main effect of logical reasoning ability on the argument generation quality.

H4.2: LRA is positively related to the argument evaluation ability.

H4.3: (a) LRA is not related to the myside bias in argument generation task; (b) the prompting of debiasing strategy does not moderate the main effect of LRA on the myside bias in argument generation task.

H4.4: LRA is not related to the myside bias in argument evaluation task.

Cognitive Reflection Test

The CRT would seem to be ideally constructed as a good predictor of performance on heuristic-and-biases tasks; however, the results have been inconsistent. Frederick (2005) observed that CRT could predict performance on temporal discounting, the tendency to choose high-expected-value gambles and framing effects. Also, Cokely and Kelley (2009) found moderate correlation between CRT and the proportion of choices consistent with expected value. Oechssler, Roider, and Schmitz (2009) found that CRT was related to the number of expected-value choices and the tendency to commit the conjunction fallacy. Koehler and James (2010) found significant correlations between CRT and the use of maximizing strategies on probabilistic prediction tasks. Toplak, West and Stanovich (2011) showed that the CRT could explain substantial

variance in rational thinking performance after partialling out cognitive ability, thinking dispositions, and executive functioning. Furthermore, they also demonstrated that the CRT actually is better predictor than intelligence, thinking dispositions and executive functioning.

On the contrary, Campiteli and Labollita (2010) found little relation between CRT and the choice of high-expected-value gambles. Obrecht, Chapman and Gelman(2009) found no relation between CRT and the degree of encountering frequency bias. Further, I found no relation between CRT and the quality of argument generation/myside bias. Therefore, I hypothesize that,

H5.1: (a) CRT is not related to the argument generation quality; (b) the prompting of debiasing strategy does not moderate the main effect of CRT on the argument generation quality.

H5.2: CRT is not related to the argument evaluation ability.

H5.3: (a) CRT is not related to the myside bias in argument generation task; (b) the prompting of debiasing strategy does not moderate the main effect of CRT on the myside bias in argument generation task.

H5.4: CRT is not related to the myside bias in argument evaluation task.

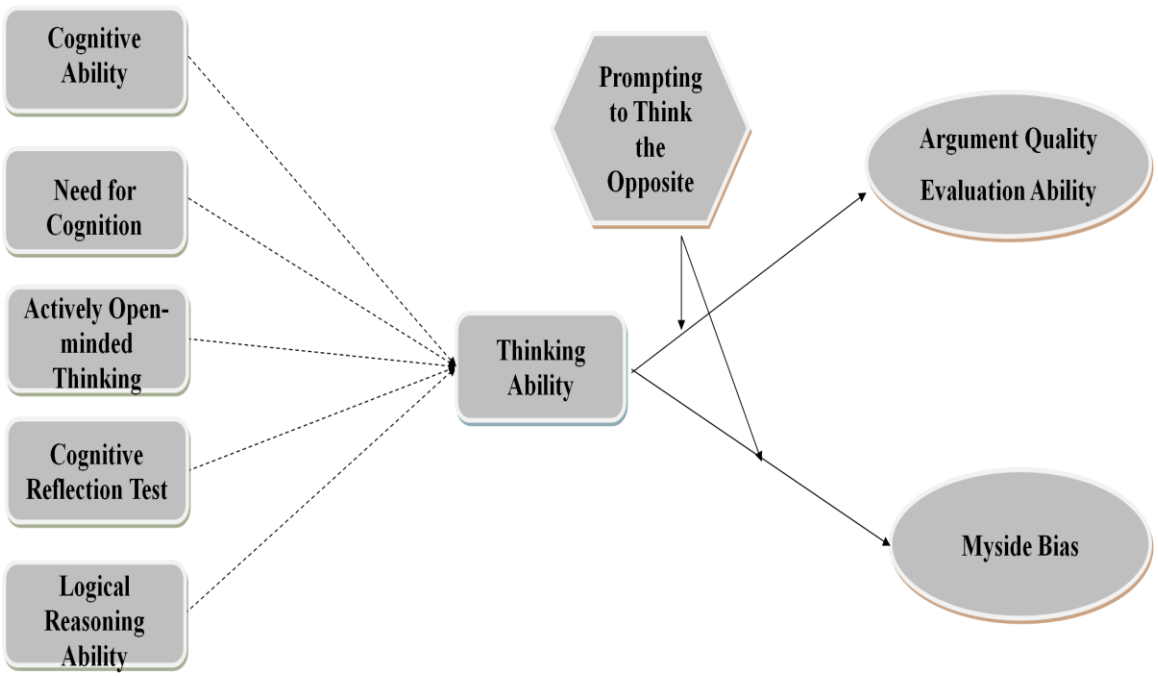

Figure 2. Study 2 Framework in Both Argument Generation and Evaluation Tasks

Methodology

Experimental Design

In order to test if the display formats of arguments in evaluation task would influence people's evaluation ability, I conducted an experiment using a 2 x 3 x 2 between design: between factors of Information Source (pros first vs. cons first), Strength Order (descending vs. ascending vs. randomized), and Display (one by one vs. all at one time) between subject factorial design experiments. Also, I needed to draw three topics out of six for each cell. Therefore, I used "Custom Design" in JMP to generate the

perfect design for the evaluation tasks. I ended up conducting 30 versions of survey and each respondent was randomly assigned into one of them.

Pretest

First, I generated the arguments for six different topics and did couple pretests to select two good and two bad arguments from both pros and cons which are differing in quality to specific topics for the focal study. Second, I asked three professors in business school to rate all 48 arguments (8 arguments for each topic; 6 topics in total) in order to make sure the quality for the same level of argument (four levels in total) in both pro and con arguments is the same. I went back and forth couple times with professors to revise all the arguments to achieve this purpose.

Procedure

Subjects comprised 283 students in a major northeastern university in the U.S. who participated in an online setting. In all cases, data was collected using a computerized Qualtrics survey. I removed 21 respondents due to the incompleteness. Therefore, 262 respondents remained in the data analyses.

First part of the survey was generation task that I used the same settings as study 1. Subjects were presented with a consumption-related issue (e.g., "Online reviews of products and services are trustworthy."; "Price equals quality: For most products, you "get what you pay for"—higher prices mean higher quality."). Each issue had a number of possible pro and con arguments associated with it. In the first beginning, subjects indicated their "spot" agreement with the statement and provided reasons for their view in an open-ended fashion. Then, all subjects were provided with successively stronger prompts to consider and list out all possible arguments about the issue, and to challenge their initial view.

The second part of the survey was two sets of evaluation task. Each respondent was asked to rate eight statements in each topic, and two topics in total. Here is the sample instruction for evaluation tasks in this study. "In the next several pages, you will see a topic presented along with a short paragraph. We are interested in how strong or weak you believe the reasoning to be. Please rate the quality of the reasoning on the scales from 0 (very bad) to 100 (very good). Use your own judgment, and try to assess the statements to the best of your ability."

After the generation and evaluation tasks, I collected a battery of individual difference measures comprising numerous categories (SAT score, Actively Open-minded Thinking, Need for Cognition, Risk Seeking, Belief in Superstitions, Conscientiousness, etc.), as well as objectively scored questions (e.g., Logical Reasoning Ability, Cognitive Reflection Test). Then, bunch of demographics information were collected.

Measurements

Dependent Variables

Overall quality/ability. Arguments that generated by respondents were independently graded for quality by one Ph.D. student in business school who is familiar with all the pros and cons of each topic and she used a consensus rubric to guide her evaluations. The argument quality score in generation task ranges from 0 to 100, meaning better quality for higher scores. In evaluation task, I had three professors rate all the 48 statements and used their ratings as rubric for evaluation quality. The score in evaluation ability is derived by the absolute difference of respondent's standardized rating and the mean of three professors' standardized rating, and then reversed by subtracting the score from the maximum score.

Myside bias. In generation task, my-side and other-side arguments were independently counted by one Ph.D. student in business school who is familiar with all the pros and cons of each topic and she used a consensus rubric to guide her evaluations. Different from study 1, the myside bias score equals to the ratio of myside argument counts to the sum of myside argument counts and otherside argument counts. In evaluation task, each participant's myside bias was indexed by a difference score calculated depending on their prior opinion. For the pro-choice group, the myside bias score was derived by adding their 2 standardized good con scores to their 2 standardized bad con scores, and subtracting from the sum of 2 standardized good pro and 2 standardized bad pro scores. For the anti-choice group, the myside bias score was derived by adding their 2 standardized good pro scores to their 2 standardized bad pro

scores, and subtracting from the sum of 2 standardized good con and 2 standardized bad con scores. Positive difference scores indicate myside bias, and the higher the score, the more the myside bias. A score of zero or negative scores may indicate unbiasedness in argument evaluation tasks or otherside bias.

Independent Variables

I adopted all the measurements of independent variables from study one.

Cognitive ability. Respondents were asked to provide their reading, mathematical, writing, total SAT scores, and how many times they took SAT in the demographics section. An indication of the validity of such self-reported scores was obtained by several studies (Stanovich & West, 1998; West & Stanovich, 1991). Also, my reliance on the total SAT score as a measure of cognitive ability is justified by the fact that it loads highly on psychometric g (Frey & Detterman, 2004), and that it is strongly associated with working memory—the quintessential indicator of computational capacity in cognitive science (Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Conway, Kane, and Engle, 2003; Engle, 2002; Engle, Tuholski, Kaughlin, & Conway, 1999; Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle, 2004). The mean reported reading SAT score of the respondents was 556 (SD_97), the mean reported mathematical SAT score was 561 (SD_94), the mean reported writing SAT score was 559 (SD_99), and the mean total SAT score was 1676 (SD_240). I treat the total SAT score as a continuous variable, indicating higher cognitive ability for higher total SAT score.

Actively open-minded thinking scale. I adopted the actively open-minded thinking scale from Stanovich and West (2007). This scale was composed of 41 items drawn from a variety of sources as follows: 8 items from the Openness-Values facet of the Revised NEO Personality Inventory (Costa & McCrae, 1992), 10 items from a flexible thinking scale developed by Stanovich and West (1997), 9 items from the belief identification scale developed by Sá, West, and Stanovich (1999), 9 items measuring dogmatism (Paulhus, & Reid, 1991; Robnison, Shaver, & Wrightsman, 1991; Troldahl, & Powell, 1965), 3 items from the categorical thinking subscale of Epstein and Meier's (1989) constructive thinking inventory, and 2 items from a counterfactual thinking scale developed by Stanovich and West (1997). All items were scored in the direction that higher scores represented a greater tendency toward open-minded thinking. Examples of items are "Certain beliefs are just too important to abandon no matter how good a case can be made against them" (reverse scored), "People should always take into consideration evidence that goes against their beliefs", "No one can talk me out of something I know is right" (reverse scored). The response format for each item in the scale was: Strongly Agree (7), Moderately Agree (6), Slightly Agree (5), Moderate (4), Slightly Disagree (3), Moderately Disagree (2), and Strongly Disagree (1). The score of AOT scale was obtained by summing the responses of the 41 items (Mean_185.81, SD_25). I treat the AOT score as a continuous variable, indicating higher actively open-minded thinking for higher AOT score.

Need for cognition scale. The need for cognition scale is an assessment instrument that quantitatively measures "the tendency for an individual to engage in and enjoy thinking" (Cacioppo & Petty, 1982). Cacioppo and Petty collaborated with Chuan

Feng Kao to shorten the scale to the 18-item format, which is popularly used in studies nowadays (Bornstein, 2004; Gauthier, Christopher, Walter, Mourad, & Marek, 2006; Coutinho & Woolery, 2004; Sadowski & Gülgös, 1992, 1996). The 18-item need for cognition scale (Cacioppo, Petty, & Kao, 1984) is adopted in this study. Sample questions are "Thinking is not my idea of fun" (reverse scored), "The notion of thinking abstractly is appealing to me", "I find satisfaction in deliberating hard and for long hours." The response format for each item in the scale was: Strongly Agree (7), Moderately Agree (6), Slightly Agree (5), Moderate (4), Slightly Disagree (3), Moderately Disagree (2), and Strongly Disagree (1). The score of NCS was obtained by summing the responses of the 18 items (Mean_78.87, SD_12.37). I treat the NCS score as a continuous variable, indicating higher thinking needs for higher NCS score.

Logical reasoning ability. I took a set of questions from logical reasoning section in LSAT and did a pretest to remove some extreme difficult and simple questions among 105 students. After pretest, I kept six questions for our study based on the results from the Item Response Analysis (IRT). Typically, the primary goal of IRT is to examine the quality and consistency of items. Also, one can use the ability estimates yielded from IRT modeling to differentiate people's ability. First, I graded the respondents' answers as correct (1) or wrong (0) and then put these scores of six questions all together into Item Response Analysis. Second, Item Response Analysis showed me the item characteristic curves (ICC), telling me the probability of answering the item correctly at different levels of respondent ability, whereas the item information function (IIF), informing me how much reliable information about the respondent I can obtain at different levels of student ability. Third, IRT gave me the respondents' ability scores based on how many questions

they answered correctly and also based on how difficult those questions were. Then, I used these ability scores as logical reasoning ability score (Mean_-0.2144, SD_1.1461) and treated it as a continuous variable.

Cognitive reflection test. The Cognitive Reflection Test (CRT) is a three-item measure introduced in the journal literature by Frederick (2005). This test is designed to measure the tendency to override a prepotent response that is incorrect and to engage in further reflection that leads to the correct response. Sample question is as follows: A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost? I graded the answers as correct (1) or wrong (0), and then added up the scores of these three questions together. I treated this score as a continuous variable, ranging from 0 to 3 (Mean_0.8626, SD_1.0340).

Results

Argument Quality

Argument Generation Task

In Table 10, Standard Least Squares Analysis showed that LRA and CRT were statistically significant ($p < 0.05$) and indicated that logical reasoning ability was the best predictors of performance on initial argument quality. Logical reasoning ability was the variable that predicted the most unique variance (6.81%). This analysis showed that logical reasoning ability is the best predictor of performance on argument quality before

prompting to think about the other-side arguments. Therefore, H1.1(a), H2.1(a), H3.1(a), H4.1(a) were supported, except H5.1(a). This means that people who are higher in logical reasoning ability or cognitive reflection test can generate better arguments to support their constructed preferences. Most importantly, I did replicate the results in study 1 and showed that cognitive ability and self-reported thinking dispositions are not related to argument generation quality, while objectively scored measures-logical reasoning ability and CRT are positively correlated with argument generation quality.

Table 11 showed that LRA and CRT overall were best predictors of all argument quality through two prompting. However, there was no interaction effect between prompting and LRA/CRT, and the positive *b* values indicated that the correlation between LRA/CRT was getting stronger after prompting. On the other hand, there was an interaction effect between prompting and NCS, showing that the relationship between argument generation quality and need for cognition changed from negative (b = -0.022) to positive (b = 0.0753 and 0.0397) after two prompting, showing that prompting people to think harder and think about the other side arguments increased the argument generation quality among people who are higher in need for cognition scale. Therefore, H1.1(b), H2.1(b), H4.1(b), H5.1(b) were not supported, which meant prompting to think the other side arguments did not change the correlation between the individual differences and argument generation quality. However, H3.1(b) was supported, which meant prompting to think the other side arguments did help people who were high in need for cognition. People who are higher in need for cognition have more argument quality increases after prompting.

Argument Evaluation Task

Based on Table 12, Standard Least Squares Analysis showed that LRA and AOT were statistically significant ($p < 0.05$) and indicated that the logical reasoning ability and actively open-minded thinking scale were best predictors of performance on argument evaluation tasks. AOT was the variable that predicted the most unique variance (8.72%) and LRA predicted the second highest unique variance (8.52%). Therefore, H1.2, H3.2, H4.2, H5.2 were supported, except H2.2. This means that people who are higher in logical reasoning ability or actively open-minded thinking scale have better argument evaluation ability. They were able to differentiate the good arguments from the bad ones.

Table 10. SLS Analysis Predicting Initial Argument Generation Quality

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	39.1345	11.5015	3.40	0.0008**
SAT	-0.0061	0.0041	-1.48	0.1394
AOT	0.0212	0.0485	0.44	0.6624
NCS	-0.0207	0.0923	-0.22	0.8227
LRA	2.7867	1.1196	2.49	0.0135**
CRT	2.5973	1.0132	2.56	0.0110**

Note: *: p<0.1; **: p<0.05

Table 11. Repeated Measures MANOVA Analysis Predicting Argument Generation Quality

Parameter	Group Main Effect		Main Effect of Prompt Within Group	
	F	*(df1, df2)*	*F*	*(df1, df2)*
Intercept	11.6729**	(1, 226)	1.0733	(2, 225)
SAT	2.9818*	(1, 226)	1.5503	(2, 225)
AOT	0.1360	(1, 226)	0.1353	(2, 225)
NCS	0.1224	(1, 226)	4.1328**	(2, 225)
LRA	5.9237**	(1, 226)	1.0484	(2, 225)
CRT	6.7721**	(1, 226)	0.4531	(2, 225)

Note: *: p<0.1; **: p<0.05

Table 12. SLS Analysis Predicting Argument Evaluation Ability

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-2.2065	1.7211	-1.28	0.2011
SAT	$-4.574e^{-5}$	0.0006	-0.07	0.9411
AOT	0.0202	0.0073	2.79	0.0057**
NCS	-0.0150	0.0138	-1.09	0.2783
LRA	0.4286	0.1675	2.56	0.0112**
CRT	0.0035	0.1516	0.02	0.9815

Note: *: p<0.1; **: p<0.05

Myside Bias

I included prior attitude strength into the myside bias model since prior belief is the biggest trigger of myside bias. So I would like to see if the individual differences were still can predict the ability of avoiding myside bias while the prior attitude strength was inside the model.

Argument Generation Task

In Table 13, Standard Least Squares Analysis showed that LRA and prior attitude strength were statistically significant ($p < 0.05$), and indicated that people who were higher in logical reasoning ability have less myside bias in argument generation task. Therefore, H1.3(a), H2.3(a), H3.3(a), and H5.3(a) were supported. However, H4.3(a) was not supported. The result showed negative relationship between logical reasoning ability and myside bias, which meant people who are higher in logical reasoning ability could suppress natural myside bias inclination and think critically by balancing out more both myside and otherside arguments to support their views instead of generating only myside arguments to win over their views. This result is different from the first study, and gives stronger evidence to prove that logical reasoning ability is a good predictor of critical thinking ability and avoiding myside bias in argument generation task.

Table 14 showed that LRA overall was the best predictor of skills in avoiding myside bias through two prompting. However, H1.3(b), H2.3(b), H3.3(b), H4.3(b), H5.3(b) were not supported, which meant prompting to think the other side arguments does not moderate the main effect of individual differences on avoiding myside bias.

Argument Evaluation Task

In Table 15, Standard Least Squares Analysis showed that CRT was statistically significant ($p < 0.05$) and indicated that CRT was the best predictor of ability to avoid myside bias in evaluation task. This means that people who are higher in CRT have less myside bias in argument evaluation task. Therefore, H1.4, H2.4, H3.4 were supported, except H4.4 and H5.4. Surprisingly, logical reasoning ability was not correlated with myside bias in argument evaluation task.

Table 13. SLS Analysis Predicting Initial Myside Bias in Argument Generation Task

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.7773	0.2129	3.65	0.0003**
SAT	$-8.513e^{-5}$	$7.932e^{-5}$	-1.07	0.2844
AOT	-0.0002	0.0009	-0.23	0.8175
NCS	0.0013	0.0017	0.78	0.4353
LRA	-0.0422	0.0211	-2.00	0.0465**
CRT	-0.0132	0.0188	-0.70	0.4828
Prior Attitude Strength	0.0050	0.0014	3.72	0.0003**

Note: *: p<0.1; **: p<0.05

Table 14. Repeated Measures MANOVA Analysis Predicting Myside Bias in Argument Generation Task

Parameter	Group Main Effect		Main Effect of Prompt Within Group	
	F	*(df1, df2)*	*F*	*(df1, df2)*
Intercept	18.3389**	(1, 208)	0.3547	(2, 207)
SAT	0.0268	(1, 208)	2.2907	(2, 207)
AOT	0.2575	(1, 208)	0.1290	(2, 207)
NCS	0.0306	(1, 208)	2.2595*	(2, 207)
LRA	3.3182*	(1, 208)	1.2532	(2, 207)
CRT	0.7890	(1, 208)	0.0289	(2, 207)
Prior Attitude Strength	9.2761**	(1, 208)	3.5369**	(2, 207)

Note: *: p<0.1; **: p<0.05

Table 15. SLS Analysis Predicting Myside Bias in Argument Evaluation Task

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-3.4760	2.6580	-1.31	0.1923
SAT	0.0007	0.0010	0.77	0.4398
AOT	0.0028	0.0112	0.25	0.8024
NCS	0.0354	0.0214	1.65	0.0995*
LRA	0.3741	0.2587	1.45	0.1496
CRT	-0.5154	0.2388	-2.16	0.0320**
Prior Attitude Strength	0.0660	0.0165	4.00	<.0001**

Note: *: p<0.1; **: p<0.05

Manipulation Check

Argument Order Effect

Table 16 showed that the order of the arguments in general did not have significant impact on individuals' performance in evaluation tasks. Therefore, the order of the arguments will not be the concern in experimental design in future studies.

Table 16. SLS Analysis Predicting Argument Order Effect

Parameter	SS	F Ratio	Prob>F
Display	2.2649	0.4367	0.5093
ProConFirst	6.7544	1.3023	0.2549
StrengthOrder	14.4103	1.3892	0.2512
Display*ProConFirst	2.7768	0.5354	0.4650
Display*StrengthOrder	0.9340	0.0900	0.9139
ProConFirst*StrengthOrder	0.3129	0.0302	0.9703
Display*ProConFirst*StrengthOrder	0.5489	0.0529	0.9485

Note: *: $p<0.1$; **: $p<0.05$

Conclusion

In this study, I reconfirmed the findings in study 1. I re-demonstrated that logical reasoning ability is the best predictor both in argument quality and ability to avoid myside bias in argument generation task. However, I found different results in argument evaluation task. Logical reasoning ability is still the good predictor of evaluation ability; however, there is no relationship between logical reasoning ability and myside bias in argument evaluation task. The differing findings show that the so-called "confirmation bias" could be different in argument generation and evaluation tasks. I will need to dig further in future studies.

Furthermore, I found that cognitive reflection test is a good predictor of the quality of argument generation task and the ability to avoid myside bias in argument evaluation task. Also, actively open-minded thinking scale is a good predictor of argument evaluation ability. Most importantly, cognitive ability is never a good predictor in both argument generation and evaluation tasks, which provide some findings to fill the gap in the literature.

Other Interesting Findings and Future Study

In Study 2, I also observed some interesting findings about attitude and confidence change differently in argument generation and evaluation tasks. Based on Table 17 and 18, the results showed that logic reasoning ability interacted with attitude

polarization and confidence increase in argument generation task. However, there was no main effect of all the individual difference factors in attitude change in evaluation task. People who were higher in logic reasoning ability had stronger opinion (attitude polarization) and higher confidence (confidence polarization) after they generated more both myside and otherside arguments. I proposed two possible mechanisms by which couple times of argument generation requests might degrade attitude and confidence for people low in logical reasoning ability, and increase attitude and confidence for people high in logical reasoning ability. The first possible explanation is that people lower in logical reasoning ability are more susceptible to the illusion of explanatory depth. People generally tend to believe they are more competent than they actually are, and then lead to overconfidence and stronger opinions in the first beginning. However, the illusion of explanatory depth arises when people overestimate their ability to explain mechanical and natural processes (Rozenblit & Keil, 2002).

The second explanation is from motivated reasoning studies. The studies of motivated reasoning show that it makes sense to look for arguments for our opinions before we find ourselves called upon to state them. People will be ready and feel confident if the search for arguments is successful. However, if the search for arguments fails, then it might be better to adopt a weaker position, one that is easier to defend. In both Study 1 and 2, the results show that people who are low in logical reasoning ability believe that they understand these concepts quite deeply and are surprised by the shallowness of their own explanations when prompted to generate the arguments to support and oppose their positions. In contrast, people who are high in logical reasoning ability build their confidence and stronger opinions by generating and weighing more

pros and cons after the prompting to consider the other side arguments (see Table 20 and Table 21).

Most interestingly, Tale 18 also showed negative relationship between confidence change and NCS/SAT. The findings indicated that people who are higher in need for cognition and cognitive ability have confidence decrease while people who are lower in need for cognition and cognitive ability have confidence increase. The plausible explanations are similar to above: people who are higher in need for cognition and cognitive ability tend to lose their confidence after they are asked to generate reasons to support their constructed preferences since they are unable to do so (see Table 20 and Table 21).

However, based on Table 19, there is no main effect in attitude change in argument evaluation task. The varied results could provide some insights for the mixed findings in attitude polarization literature. I will discuss this issue deeper in Study 3.

In attitude change and group polarization literature, there are mixed findings. According to the Argumentative Theory, people are all skilled arguers. People will be good at reasoning once they are in argumentative settings, and will lead to good solutions and depolarization. However, based on the findings of Study 1 and 2, I propose that the moderating effect of individual difference can provide some explanations to this stream of literature. I believe that not everyone is good thinker, even after using the debiasing strategy (e.g., prompting to consider the opposite side) (Kuhn, 1992) and arguing with other people won't always end up with good outcomes and depolarization. I will investigate this issue further in Study 3.

Table 17. SLS Analysis Predicting Attitude Polarization in Argument Generation Task

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-6.4143	11.7903	-0.54	0.5870
SAT	0.0045	0.0042	1.07	0.2868
AOT	0.0153	0.0497	0.31	0.7578
NCS	-0.1191	0.0946	-1.26	0.2093
LRA	2.7312	1.1477	2.38	0.0182**
CRT	-0.6364	1.0387	-0.61	0.5407

Note: *: p<0.1; **: p<0.05

Table 18. SLS Analysis Predicting Confidence Change in Argument Generation Task

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	47.9541	13.9692	3.43	0.0007**
SAT	-0.0111	0.0050	-2.20	0.0285**
AOT	-0.0733	0.0589	-1.25	0.2142
NCS	-0.2325	0.1120	-2.07	0.0391**
LRA	3.7180	1.3598	2.73	0.0067**
CRT	-0.4574	1.2306	-0.37	0.7105

Note: *: p<0.1; **: p<0.05

Table 19. SLS Analysis Predicting Attitude Polarization in Argument Evaluation Task

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	3.2977	11.6537	0.28	0.7775
SAT	-0.0022	0.0042	-0.52	0.6028
AOT	-0.0452	0.0491	-0.92	0.3582
NCS	0.0425	0.0935	0.46	0.6495
LRA	1.2666	1.1344	1.12	0.2654
CRT	0.7484	1.0266	0.73	0.4668

Note: *: p<0.1; **: p<0.05

Table 20. SLS Analysis Predicting Relevant Counts in Argument Generation Task

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	3.4928	1.0540	3.31	0.0011**
SAT	-0.0006	0.0004	-1.53	0.1272
AOT	-0.0005	0.0044	-0.10	0.9167
NCS	-0.0056	0.0085	-0.67	0.5053
LRA	0.2514	0.1026	2.45	0.0150**
CRT	0.2096	0.0929	2.26	0.0250**

Note: *: p<0.1; **: p<0.05

Table 21. SLS Analysis Predicting Otherside Counts in Argument Generation Task

Parameter	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.5976	0.5420	1.10	0.2714
SAT	$4.405e^{-6}$	0.0002	0.02	0.9820
AOT	0.0010	0.0023	0.42	0.6755
NCS	-0.0046	0.0043	-1.06	0.2895
LRA	0.1331	0.0528	2.52	0.0124**
CRT	0.0603	0.0478	1.26	0.2081

Note: *: $p<0.1$; **: $p<0.05$

CHAPTER 5

STUDY THREE

Introduction

Most consumption decisions will not just be made by oneself, but with families, friends or colleagues, especially the complicated decisions or the decisions with higher financial risks. People frequently ask for advices and suggestions, or sometimes argue with each other to find out which choice is the best. Since reaching the optimal decisions is the goal for all consumers, I will investigate the best way to get decisions done. However, in Study 1 and 2, I found that not everyone is good at making decisions by themselves. People are different in logical reasoning ability, which means some people tend to think critically and try to find the best solutions by weighing pros and cons, while some people are susceptible to confirmation bias and one-sided thinking, and try to justify themselves and then lead to suboptimal decisions (Kuhn, 1992; Nickerson, 1998; Mercier & Sperber, 2011).

Information can lead to an increase in confidence and may or may not have an effect on accuracy (Davis, Lohse, & Kottemann, 1994; Oskamp, 1965; Paese & Sniezek, 1991; Peterson & Fitz, 1988). Lord, Ross, and Leper (1979) showed a violation of the neutral-evidence principle in a situation where such errors in daily life have deadly serious consequences. It seems that using irrelevant information to boost oneself's confidence in his/her decisions is not wise and may lead to suboptimal decisions.

Most research attributes this behavior to confirmation bias, which proposes that people tend to distort information to support their own beliefs. According to the notorious confirmation bias theory, people are inclined to their favored choice by searching, interpreting, memorizing the information biased to their preferences.

The other explanation could be that people update their belief by using Bayes' inference, a set of procedures based upon Bayes' theorem and the subjectivist interpretation of probability. These procedures show how to (a) identify the data sources that are most useful for discriminating between competing hypotheses, (b) assess the implications of an observed datum in terms of the truth of competing hypotheses, (c) aggregate different data implications into an overall appraisal of the relative likelihood of those hypotheses being correct, (d) use that appraisal to select the course of action that seems best in light of available evidence (Fischhoff & Beyth-Marom, 1983). Bayesians have more frequent recourse to the theorem because the subjectivist position enables them to incorporate prior beliefs explicitly in their inferential processes.

Since people update the data by incorporating prior beliefs, people with lower logical reasoning ability will be hard to decouple from prior belief and be prone to distort the new information toward their favored stance. Furthermore, it may be possible that people just assume that there are fewer opposing arguments after seeing more irrelevant arguments. In this situation, people are not suffering from confirmation bias, but just update their belief by eliminating more irrelevant information. Both situations will lead to more confidence in their beliefs; however, they are initiated by different causes.

Following normative neutral evidence principle, people's beliefs should not be influenced by neutral evidence and then lead to increase in confidence (Baron, 2000). Neutral evidence refers to evidence that is equally consistent with a belief and its converse. Therefore, neutral evidence might consist of mixed evidence with equal weights, ambiguous evidence, or irrelevant evidence.

Confidence is determined through a hypothesis testing process (Peterson & Fitz, 1988). When people express their confidence in a previously constructed statement, it is assumed that the level of confidence is determined by salient factors that they believe ought to influence the accuracy of the statement. However, people often believe that irrelevant variables should influence their accuracy (Kahneman, Slovic, & Tversky, 1982), and then it supposes to covary with their confidence too. The amount of information is usually a salient variable that may be expected to affect confidence, no matter if the information is related to the topic or not since sometimes people are unable to recognize when information is of no value (Oskamp, 1965; Peterson & Fitz, 1988).

However, in the first study, I showed that not everyone suffered from this bias, but only people with low logical reasoning ability. The results showed that people with high logical reasoning ability would be able to overcome the confirmation bias and think critically by generating both-sides arguments. These people are able to build the confidence and have stronger opinions by generating and weighing both-sides arguments. I also found that people with low logical reasoning ability had less confidence and weaker attitudes after they were asked to generate the arguments to support their statements and constructed preferences.

In the literature, several studies have showed that the inclusion of irrelevant information may not increase prediction accuracy, but may increase confidence. In this study, I do not ask respondents to predict, but ask them to evaluate arguments and to see how they update their prior beliefs. Also, in order to compare the results in argument generation and evaluation tasks, each respondent took one set of argument generation task and one set of evaluation task.

In this study, I investigated the moderating effect of individual differences on six debiasing strategies: argument generation (considering the opposite side, group think, argumentation settings), and argument evaluation (agreement evaluation, disagreement evaluation, half agreement & half disagreement evaluation settings). I would like to find out what kind of debiasing strategy is better for people in different levels of logical reasoning ability and other individual differences. Also, I investigated how these strategies change the attitude and confidence in people with different levels of logical reasoning ability and other individual differences. Biased search, interpretation and memory contribute to overconfidence, attitude polarization in the face of the contrary evidence, belief perseverance, the irrational primacy effect and illusory correlation. Further, poor thinking could be caused by too little search, by overconfidence, and by biases in favor of possibilities that are favored initially. On the contrary, good thinking consists of (1) thorough search based on the importance of the question, (2) appropriate confidence in the quality of thinking done, and (3) fairness to all the possibilities. As being said, pushing people away from overconfidence and extreme opinion will lead to the path of good thinking. Therefore, consistent with the literature, I see attitude and confidence moderation are good signs of avoiding poor thinking. I investigated and

compared the effects of six debiasing strategies on attitude polarization and confidence change in this study. Further, I discussed the moderating effects of individual differences on these debiasing strategies.

Literature Review

Confirmation Bias

One of the most studied biases in psychology, the notorious confirmation bias consists in the "seeking, interpreting, memorizing evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand" (Nickerson, 1998). It seems that everybody is affected by confirmation bias to some degree irrespective of factors like general intelligence or open mindedness (Klaczynski, Gordon, & Fauth, 1997; Klaczynski & Lavallee, 2005; Stanovich & West, 2007, 2008). In Study 1 and 2, I also demonstrated that people were prone to generate supporting arguments to their opinions instead of counterarguments. However, individual difference has some variations in the debiasing strategies (e.g., consider the opposite side, or decoupling from prior belief) (Stanovich & West, 1997, 1998, 2000, 2007, 2008). In Study 1 and 2, I also found that people with higher logical reasoning ability would be able to produce more counterarguments and have less myside bias. In Study 2, I reconfirmed the findings of argument generation task in Study 1. However, I found the differing effects of individual differences, especially logical reasoning ability, on ability of avoiding myside bias in argument generation and evaluation tasks. Therefore, I project that the individual

difference may have varying effects on argument generation and evaluation tasks. In Study 3, I look further into this issue.

Argumentative Theory

For standard theories of reasoning, the confirmation bias is no more than a flaw of reasoning, however, for the argumentative theory, the confirmation bias is a consequence of the function of reasoning and hence a feature of reasoning when used for the argument production (Mercier & Sperber, 2011). The argumentative theory puts such well-known demonstrations of "irrationality" in a novel perspective. They propose that human reasoning is a remarkably efficient specialized device adapted to a certain type of social and cognitive interaction at which it excels, instead of a profoundly flawed general mechanism.

According to the Argumentative Theory (Mercier & Sperber, 2011), when one is alone (individual) or with people who hold similar views (group think), one's arguments will not be critically evaluated or challenged. This is when the confirmation bias is most likely to lead to poor outcomes and suboptimal decisions. However, when reasoning is used in a more argumentative context—that is, in arguments among people who disagree with each other but have a common interest in the truth—the confirmation bias contributes to an efficient form of division of cognitive labor. In that situation, a group can work more efficiently to solve a problem if each individual looks mostly for

arguments supporting a given solution, and brings these arguments to be tested by other group members.

The argumentative theory also proposes that the "confirmation bias" should be active only in the argument generation but not in argument evaluation tasks, no matter in discussions or group performance. They argue that people tend to be more objective in argument evaluation than in argument generation (Mercier & Sperber, 2011). They believe that the main function of reasoning is argumentative, and that's the reason why the past research in reasoning can't find good outcomes from abstract reasoning tasks. Therefore, people improve their argument quality significantly when they have to debate with other people (Blanchette & Dunbar, 2001; Kuhn, Shaw, & Felton, 1997; Mercier & Sperber, 2011).

Further, they suggest that the joint dialogic approach is much more efficient than one that each individual has to examine all possible solutions carefully on his or her own. The advantages of the confirmation bias are even more obvious given that each participant in the group discussion is in a better position to look for arguments in favor of one's favored choice (situations of asymmetrical information). Further, some studies show that the teaching of critical thinking skills to overcome the bias on a purely individual basis does not seem to yield very good results (Ritchart & Perkins, 2005; Willingham, 2008). However, in the first two studies, I found that people with higher logical reasoning ability does overcome the bias after the prompting to generate/consider otherside arguments. Also, even Mercier and Sperber (2011) mentioned that some individuals may develop some limited ability to distance themselves from their own opinion, to consider alternatives and thereby become more objective.

Attitude Polarization and Confidence Change

When people with opposing views interpret new information in a biased way, their views can move even further apart. This is called "attitude polarization". People will hold more extreme positions after they are exposed to the opposite side statements since people are motivated to reason back. Motivated reasoning leads to a biased assessment: Arguments with unfavored conclusions are rated as less sound and less persuasive than arguments with favored conclusions (Edwards & Smith, 1996; Eagly, Kulesa, Brannon, Shaw, & Hutson-Comeaux, 2000). Therefore, according to motivated reasoning and confirmation bias, people will lead to polarization both in attitude and confidence after they are exposed to otherside arguments since they tend to comb through these arguments for flaws and end up finding some. Therefore, attitude polarization is most easily observed in participants who are most knowledgeable (Braman, 2009; Redlawsk, 2002; Taber & Lodge, 2006). On reading an argument with a counterattitudinal conclusions, more knowledgeable people may find so many flaws and counterarguments to these counterattitudinal arguments, leading to attitude polarization (Lord, Ross, & Lepper, 1979; Greenwald , 1969; Pomerantz, Chaiken, & Tordesillas, 1995). This could provide some explanations to the findings in the first study: people with higher logical reasoning ability lead to attitude polarization after generating more both myside and otherside arguments, while people with lower logical reasoning ability lead to attitude depolarization after being asked to generate more myside and otherside arguments since they were not able to do that.

Motivated reasoning also affects confidence. When people construct their positions, they will be spontaneously tempted to generate reasons supporting their

postions. This may then cause them to be overconfident in the answer (Koriat, Lichtenstein, & Fischhoff, 1980).

On the contrary, the Argumentative Theory is tied up with the framework of Persuasive Argument Theory. The basic idea of persuasive argument theory is that people's preferences will be moved by the difference of the total number of persuasive pros and cons. Therefore, if you are arguing with the other person over one topic, and the otherside arguments are persuasive and original, your attitude will move to closer to the opposite side. This phenomenon is called attitude depolarization. The argumentative theory proposes that when group members disagree, discussion often lead to depolarizatoin (Kogan & Wallach, 1966; Vinokur & Burnstein, 1978).

Based on the literature and the findings in Study 1 and 2, I hypothesize that

H1: Individual and group think settings will lead to attitude polarization, while argumentative setting will lead to attitude moderation.

H2: Agreement evaluation setting will lead to attitude polarization, while disagreement and half agreement & half disagreement settings will lead to attitude moderation.

H3: Individual and group think settings will lead to confidence increase, while argumentative setting will lead to confidence decrease.

H4: Agreement evaluation setting will lead to confidence increase, while disagreement and half agreement & half disagreement settings will lead to confidence decrease.

However, I believe that this is not always the case. Individual differences should play a role in attitude and confidence change in different settings.

H5: LRA will have moderating effect in H1 to H4.

H6: CRT will have moderating effect in H1 to H4.

H7: AOT will have moderating effect in H1 to H4.

H8: NCS will have moderating effect in H1 to H4.

Methodology

Experimental Design

I conducted an experiment using a 3 x 3 between design: between factors of Argument Evaluation Task (agreement, disagreement, half agreement & half disagreement), and Argument Generation Task (individual prompting vs. group think setting vs. argumentative setting) between subject factorial design experiments. Also, I needed to draw two topics out of six for each cell. Therefore, I used "Custom Design" in JMP to generate the perfect design for this study. I ended up conducting eighteen versions of survey and each respondent was randomly assigned into one of them. I used the statements from study two to compose the group discussion in both argument evaluation and argument generation tasks.

Procedure

I recruited one hundred ten English speaking respondents using Amazon's Mechanical Turk and participated in return for couple dollars. Then, I removed sixteen incompletes and remained ninety-four completes. Participants were 50% male and 50% female, with an average age of 33.3 years. Participants' reported completed highest education were 8.8% High School Graduate or the equivalent, 23.24% Some College Credits without Degree, 1.1% Trade/Technical/Vocational Training, 9.10% Associate Degree, 45.48% Bachelor's Degree, and 8.8% Master/Professional Degree.

Participants were randomly assigned to eighteen different cells. First part of the survey is evaluation task. Subjects were presented with a consumption-related issue (e.g., "Online reviews of products and services are trustworthy."; "Price equals quality: For most products, you "get what you pay for"—higher prices mean higher quality."). Each issue had a number of possible pro and con arguments associated with it. Respondents were asked to indicate their initial views, confidence, understanding and interest in one of the six issues. Then, they read through one of the three group discussion settings. In agreement and disagreement settings, I used the display logic to present the group discussion based on their initial view about the issue. In agreement setting, subjects were presented a group discussion, containing four statements in the same view about the issue as them. In disagreement setting, subjects were presented a group discussion, containing four statements in the opposite view from them. Also, they were asked to check the statements that they were new to them and the ones that they agreed with. In half agreement and half disagreement setting, subjects were presented a group debate, no

matter which view they had. After the first manipulation task, respondents indicated their posterior view, confidence, and understanding about this issue.

Second part of the survey is argument generation task. In individual prompting setting, subjects indicated their "spot" agreement with the statement and provided reasons for their view in an open-ended fashion. Then, all subjects were provided with successively stronger prompts to consider all possible views about the argument, and to challenge their initial view. In group think setting, respondents were exposed to the same side arguments couple times, making them feel like they are in a group discussion in an online forum, and then were asked to indicate their posterior attitude, confidence and understanding. In argumentative setting, I challenged respondents the other side arguments four times, making them feel like they are arguing with other people in a group discussion in an online forum, and asked them to counter-argue with them, write down their final arguments and indicate their attitude, confidence and understanding about the issue again.

After the argument evaluation and argument generation tasks, I collected a battery of individual difference measures comprising numerous categories (Actively Open-minded Thinking, Need for Cognition, etc.), as well as objectively scored questions (e.g., Logical Reasoning Ability, Cognitive Reflection Test). Then, bunch of demographics information were collected.

Measurements

Dependent Variables

Dependent variables include, attitude before and after evaluation/generation task (responses were made using 101-point sliders from 0, strongly disagree, to 100, strongly agree) , confidence before and after evaluation/generation task (responses were made using 101-point sliders from 0, not confident at all, to 100, very confident), attitude polarization (if the respondent's initial position was equal to or larger than 50, then his/her attitude polarization score equaled to post rating minus prior rating; if the respondent's initial position was smaller than 50, then his/her attitude polarization score equaled to prior rating minus post rating). I transformed raw ratings of attitudes on issues into a measure of attitude extremity by subtracting the midpoint of the slider (50) and taking the absolute value.

Independent Variables

Independent variables include evaluation task (agreement, disagreement, half agreement & half disagreement), generation task (individual prompting vs. group think setting vs. argumentative setting), logical reasoning ability, cognitive reflection test, need for cognition, and actively open-minded thinking scale. All the individual differences measurements are the same as Study 1 and 2.

Results

Attitude Polarization

Argument Generation Task

I conducted a repeated measures ANOVA with timing of judgment (pre-generation vs. post-generation) as within-subject factor. I predicted that positions would become more moderate following argumentative task, however, become more extreme following group think and individual prompting settings. This prediction was confirmed (see Figure 3), with the interaction effect of judgment timing and generation task significant (Argumentative Setting: pre-generation-rating condition: M = 25.37, SE = 16.19, post-generation-rating condition: M = 22.13, SE = 14.59; Group Think Setting: pre-generation-rating condition: M = 27.06, SE = 16.64, post-generation-rating condition: M = 30, SE = 14.91; Individual Prompting Setting: pre-generation-rating condition: M = 25.59, SE = 16.18, post-generation-rating condition: M = 28.03, SE = 14.22), $F(2, 91) = 3.13$, $p = .0486$. There, H1 was supported.

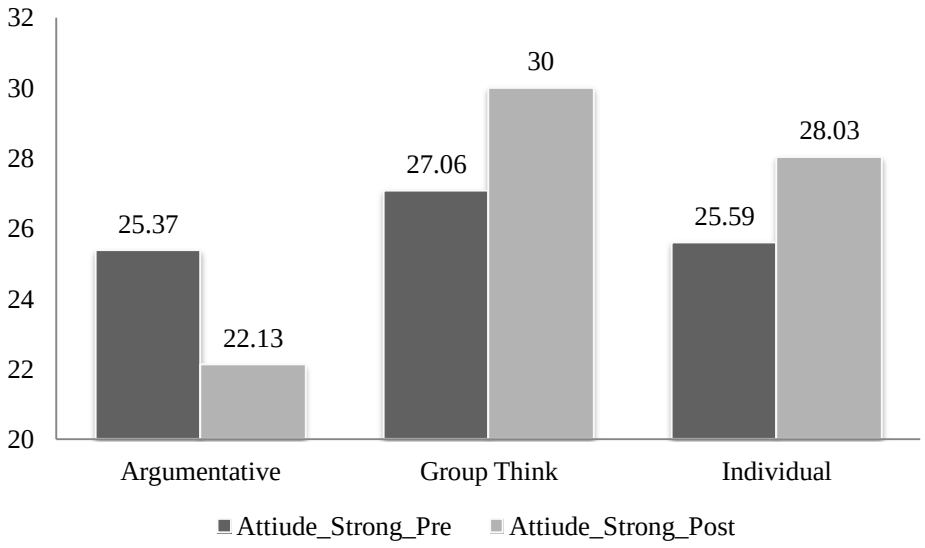

Figure 3. Interaction Effect of Timing and Argument Generation Task on Attitude Polarization

Logical reasoning ability. A regression was performed on attitude polarization with independent variables (i) logical reasoning ability, (ii) a categorical variable-generation task, and (iii) their interaction. The results showed no main effect of Logical Reasoning Ability ($F(1, 93) = 0.24$, $p = .63$), and no two-way interaction between logical reasoning ability and the generation task ($F(2, 92) = 0.0028$, $p = .9972$). Therefore, H5.1 was not supported.

Cognitive reflection test. A regression was performed on attitude polarization with independent variables (i) CRT, (ii) a categorical variable-generation task, and (iii) their interaction. The results showed a marginally significant main effect of CRT ($F(1, 93) = 3.61$, $p = .06$) , but no two-way interaction between CRT and the generation task ($F(2, 92) = 0.50$, $p = .61$). The main effect of CRT ($b = 2.11$) showed that participants

with higher CRT had more opinion polarization across all three generation tasks. However, H6.1 was not supported.

Actively open-minded thinking. A regression was performed on attitude polarization with independent variables (i) AOT, (ii) a categorical variable-generation task, and (iii) their interaction. The results showed no main effect of AOT ($F(1, 93) = 1.85$, $p = .18$), and no two-way interaction between AOT and the generation task ($F(2, 92) = 0.84$, $p = .43$). Therefore, H7.1 was not supported.

Need for cognition. A regression was performed on attitude polarization with independent variables (i) NCS, (ii) a categorical variable-generation task, and (iii) their interaction. The results showed no main effect of NCS ($F(1, 93) = 0.48$, $p = .49$), and no two-way interaction between need for cognition and the generation task ($F(2, 92) = 0.23$, $p = .79$). Therefore, H8.1 was not supported.

Argument Evaluation Task

I conducted a repeated measures ANOVA with timing of judgment (pre-evaluation vs. post-evaluation) as within-subject factor. I predicted that positions would become more moderate following disagreement and half agreement & half disagreement tasks, however, become more extreme following agreement setting. This prediction was confirmed (see Figure 4), with the interaction effect of judgment timing and evaluation task significant (Agreement Setting: pre-evaluation-rating condition: M = 25.61, SE = 14.63, post-evaluation-rating condition: M = 29.10, SE = 14.17; Disagreement Setting: pre-evaluation-rating condition: M = 33.5, SE = 13.11, post-evaluation-rating condition:

M = 28.37, SE = 15.81; Half & Half Setting: pre-evaluation-rating condition: M = 27.76, SE = 12.63, post-evaluation-rating condition: M = 24.94, SE = 14.78), $F(2, 91) = 5.62$, $p = .0050$. Therefore, H2 was supported.

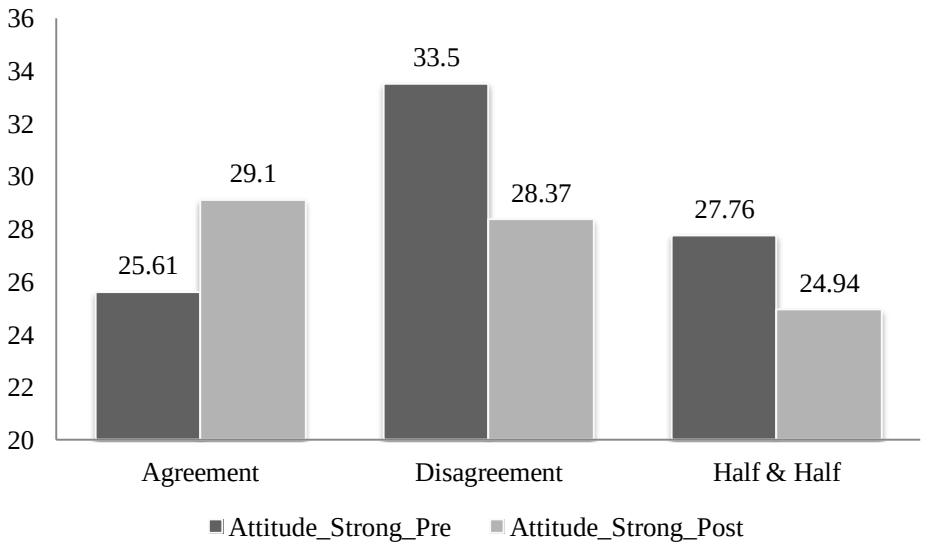

Figure 4. Interaction Effect of Timing and Argument Evaluation Task on Attitude Polarization

Logical reasoning ability. A regression was performed on attitude polarization with independent variables (i) logical reasoning ability, (ii) a categorical variable-argument evaluation task, and (iii) their interaction. The results showed strong main effect of argument evaluation task ($F(2,92) = 6.43$, $p = .0025$), no main effect of Logical Reasoning Ability ($F(1, 93) = 0.34$, $p = .56$), and a significant two-way interaction between logical reasoning ability and the evaluation task ($F(2, 92) = 3.84$, $p = .03$). To explore the interaction, I examined the slopes of logical reasoning ability at each level of argument evaluation tasks. The slope of logical reasoning ability was significant and

positive (b = 2.47) in Agreement Evaluation Setting, while the slope of logical reasoning ability was significant and negative (b = -3.91) in Disagreement Evaluation Setting, and the slope of logical reasoning ability was significant and negative (b = -0.27) in Half & Half Evaluation Setting. In addition, a spotlight analysis at one standard deviation above the mean of logical reasoning ability showed a significant difference such that high logical reasoning ability respondents had stronger attitude polarization in agreement evaluation setting versus in disagreement and half & half evaluation settings. The results showed that in agreement evaluation setting, people with higher logical reasoning ability had stronger attitude polarization, while in disagreement and half & half evaluation settings, people with higher logical reasoning ability had stronger attitude depolarization. Therefore, people with higher logical reasoning ability will be better off by reading the arguments from the opposite views. Therefore, H5.2 was supported.

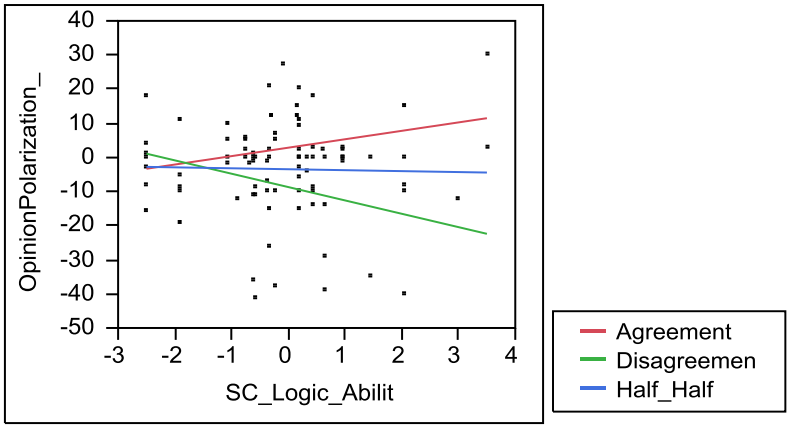

Figure 5. Interaction Effect of Logical Reasoning Ability and Argument Evaluation Task on Attitude Polarization

Cognitive reflection test. A regression was performed on attitude polarization with independent variables (i) CRT, (ii) a categorical variable-evaluation task, and (iii) their interaction. The results showed significant main effect of evaluation task ($F(2, 92) = 4.86$, $p = 0.0099$), no main effect of CRT ($F(1, 93) = 0.92$, $p = .36$), and a marginally significant two-way interaction between CRT and the evaluation task ($F(2, 92) = 2.76$, $p = .06$). To explore the interaction, I examined the slopes of CRT at each level of evaluation tasks. The slope of CRT was significant and positive ($b = 3.47$) in Agreement Evaluation Setting and the slope of CRT was significant and negative ($b = 1.64$) in Half & Half Evaluation Setting, while the slope of CRT was significant and negative ($b = -2.38$) in Disagreement Evaluation Setting. In addition, a spotlight analysis at one standard deviation above the mean of CRT showed a significant difference such that high CRT respondents had stronger attitude polarization in agreement evaluation setting and half & half versus in disagreement evaluation settings. The results showed that in agreement and half & half evaluation settings, people with higher CRT had stronger attitude polarization, while in disagreement evaluation setting, people with higher CRT had stronger attitude depolarization. Therefore, people with higher CRT will be better off by reading the arguments from the opposite views. Therefore, H6.2 was supported.

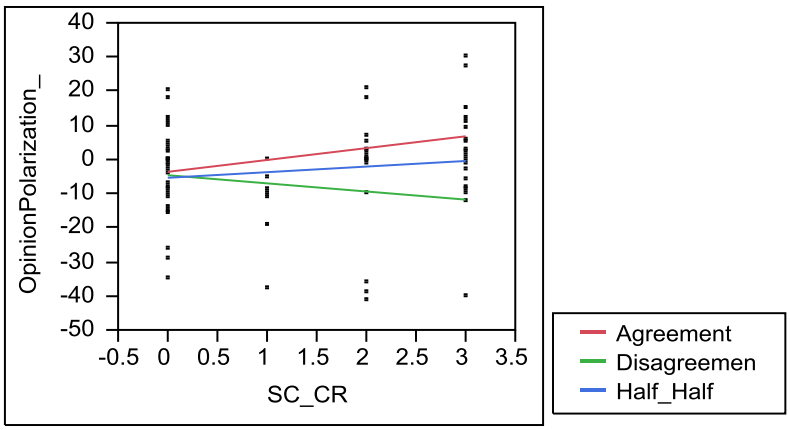

Figure 6. Interaction Effect of Cognitive Reflection Test and Argument Evaluation Task on Attitude Polarization

Actively open-minded thinking. A regression was performed on attitude polarization with independent variables (i) AOT, (ii) a categorical variable-evaluation task, and (iii) their interaction. The results showed no main effect of AOT ($F(1, 93) = 0.08$, $p = .77$), and no two-way interaction between AOT and the evaluation task ($F(2, 92) = 2.26$, $p = .11$). Therefore, H7.2 was not supported.

Need for cognition. A regression was performed on attitude polarization with independent variables (i) NCS, (ii) a categorical variable-evaluation task, and (iii) their interaction. The results showed no main effect of NCS ($F(1, 93) = 0.24$, $p = .63$), and no two-way interaction between need for cognition and the evaluation task ($F(2, 92) = 0.33$, $p = .72$). Therefore, H8.2 was not supported.

Confidence Change

Argument Generation Task

I conducted a repeated measures ANOVA with timing of judgment (pre-generation vs. post-generation) as within-subject factor. I predicted that respondents would become less confident following argumentative setting; however, they would become more confident following group think and individual prompting settings. This prediction was not confirmed (see Figure 5), without the interaction effect of judgment timing and generation task significant (Argumentative Setting: pre-generation-rating condition: M = 76.73, SE = 19.85, post-generation-rating condition: M = 75.3, SE = 15.98; Group Think Setting: pre-generation-rating condition: M = 80.38, SE = 16.27, post-generation-rating condition: M = 77.97, SE = 23.30; Individual Prompting Setting: pre-generation-rating condition: M = 80.47, SE = 17.59, post-generation-rating condition: M = 85.53, SE = 12.15), $F(2, 91) = 1.91$, $p = .1541$. Interestingly, group discussion, no matter argumentative or group think settings, lowers respondents' confidence; while individual generation task increase respondents' confidence. Therefore, H3 was not supported.

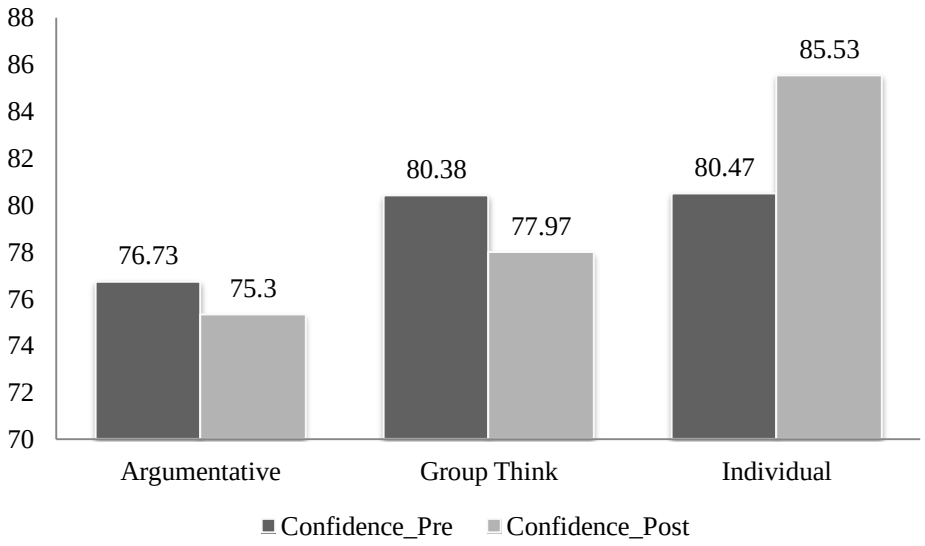

Figure 7. Interaction Effect of Timing and Argument Generation Task on Confidence Change

Logical reasoning ability. A regression was performed on confidence change with independent variables (i) logical reasoning ability, (ii) a categorical variable-argument generation task, and (iii) their interaction. The results showed no main effect of Logical Reasoning Ability ($F(1, 93) = 0.02$, $p = .88$), and no two-way interaction between logical reasoning ability and argument generation task ($F(2, 92) = 0.79$, $p = .46$). Therefore, H5.3 was not supported.

Cognitive reflection test. A regression was performed on confidence change with independent variables (i) CRT, (ii) a categorical variable-argument generation task, and (iii) their interaction. The results showed a marginally significant main effect of argument generation task ($F = 2.84$, $p = .06$), a significant main effect of CRT ($F(1, 93) = 7.12$, $p = .0091$) , and a marginally significant two-way interaction between CRT and the

argument generation task ($F(2, 92) = 2.56$, $p = .08$). The main effect of CRT ($b = 3.39$) showed that participants with higher CRT had more confidence increase after all three argument generation tasks. To explore the interaction, I examined the slopes of CRT at each level of argument generation tasks. The slope of CRT was significant and positive ($b = 6.73$) in Group Think Setting and the slope of CRT was significant and positive ($b = 3.70$) in Individual Prompting Setting, while the slope of CRT was significant and negative ($b = -0.26$) in Argumentative Setting. In addition, a spotlight analysis at one standard deviation above the mean of CRT showed a significant difference such that high CRT respondents had more confidence increase in group think and individual prompting settings versus in argumentative setting. The results showed that in group think and individual prompting settings, people with higher CRT had stronger confidence increase; while in argumentative setting, people with higher CRT had stronger confidence decrease. Therefore, people with higher CRT will be better off by arguing with people having different opinions. Therefore, H6.3 was supported.

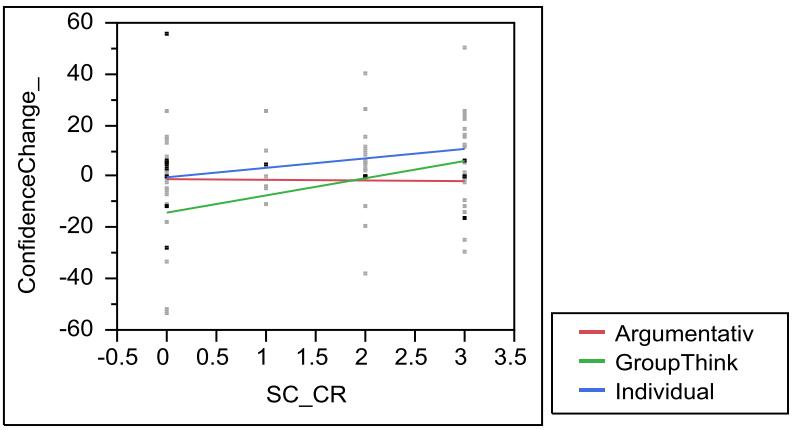

Figure 8. Interaction Effect of Cognitive Reflection Test and Argument Generation Task on Confidence Change

Actively open-minded thinking. A regression was performed on confidence change with independent variables (i) AOT, (ii) a categorical variable-argument generation task, and (iii) their interaction. The results showed no main effect of AOT ($F(1, 93) = 1.50$, $p = .22$), and no two-way interaction between AOT and argument generation task ($F(2, 92) = 0.01$, $p = .99$). Therefore, H7.3 was not supported.

Need for cognition. A regression was performed on confidence change with independent variables (i) NCS, (ii) a categorical variable-argument generation task, and (iii) their interaction. The results showed no main effect of NCS ($F(1, 93) = 0.85$, $p = .36$) , and a significant two-way interaction between NCS and argument generation task ($F(2, 92) = 4.47$, $p = .01$). To explore the interaction, I examined the slopes of NCS at each level of generation tasks. The slope of NCS was significant and positive ($b = 0.49$) in Group Think Setting and the slope of NCS was negative ($b = -0.22$) in Individual Prompting Setting, while the slope of NCS was 0 ($b = 0$) in Argumentative Setting. In

addition, a spotlight analysis at one standard deviation above the mean of NCS showed a significant difference such that high NCS respondents had more confidence increase in group think setting versus in argumentative and individual prompting settings. The results showed that in group think setting, people with higher NCS had stronger confidence increase; while in individual prompting setting, people with higher NCS had less confidence increase. Therefore, people with higher NCS will be better off by arguing with people having different opinions or think harder by themselves. Therefore, H8.3 was supported.

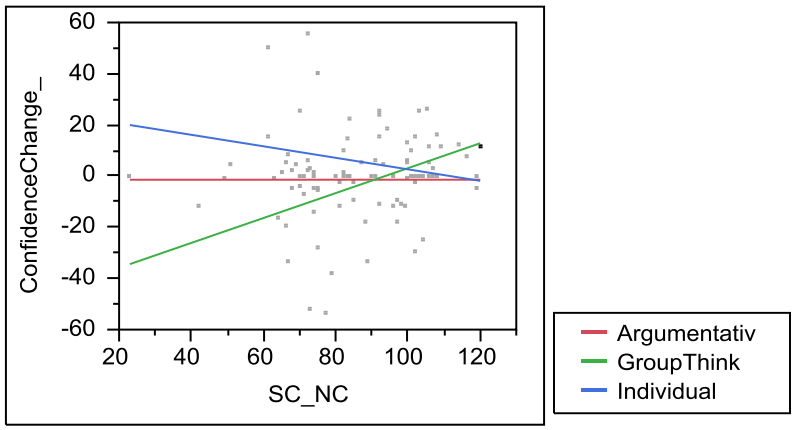

Figure 9. Interaction Effect of Need for Cognition and Argument Generation Task on Confidence Change

Argument Evaluation Task

I conducted a repeated measures ANOVA with timing of judgment (pre-evaluation vs. post-evaluation) as within-subject factor. I predicted that confidence

would become less following disagreement and half agreement & half disagreement tasks, however, become more following agreement setting. This prediction was not confirmed (see Figure 6), without the interaction effect of judgment timing and evaluation task significant (Agreement Setting: pre-evaluation-rating condition: M = 84.39, SE = 14.02, post-evaluation-rating condition: M = 88.61, SE = 14.24; Disagreement Setting: pre-evaluation-rating condition: M = 83.8, SE = 15.51, post-evaluation-rating condition: M = 80.6, SE = 19.98; Half & Half Setting: pre-evaluation-rating condition: M = 82.18, SE = 14.15, post-evaluation-rating condition: M = 84.03, SE = 15.77), $F(2, 91) = 1.43$, $p = .2451$. Interestingly, agreement evaluation task increases respondents' confidence; while disagreement evaluation task lowers respondents' confidence. However, H4 was not supported.

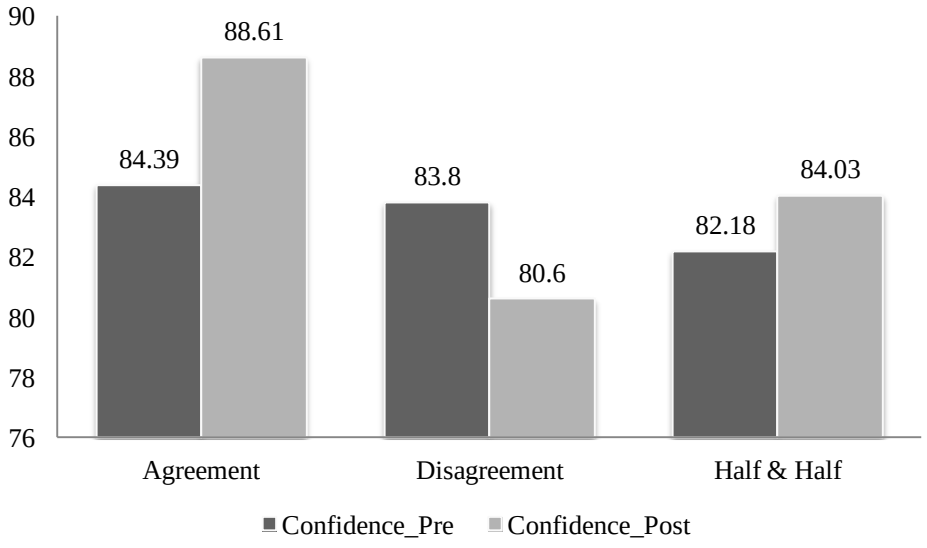

Figure 10. Interaction Effect of Timing and Argument Evaluation Task on Confidence Change

Logical reasoning ability. A regression was performed on confidence change with independent variables (i) logical reasoning ability, (ii) a categorical variable-argument evaluation task, and (iii) their interaction. The results showed a significant main effect of logical reasoning ability ($F(1, 93) = 4.48$, $p = .04$) , but no two-way interaction between logical reasoning ability and argument generation task ($F(2, 92) = 0.98$, $p = .38$). The main effect of logical reasoning ability ($b = 2.97$) showed that participants with higher logical reasoning ability had more confidence increase across all three argument evaluation tasks. However, H5.4 was not supported.

Cognitive reflection test. A regression was performed on confidence change with independent variables (i) CRT, (ii) a categorical variable-argument evaluation task, and (iii) their interaction. The results showed no main effect of CRT ($F(1, 93) = 0.0082$, $p = .93$), and no two-way interaction between CRT and argument evaluation task ($F(2, 92) = 0.51$, $p = .60$). Therefore, H6.4 was not supported.

Actively open-minded thinking. A regression was performed on confidence change with independent variables (i) AOT, (ii) a categorical variable-argument evaluation task, and (iii) their interaction. The results showed no main effect of AOT ($F(1, 93) = 0.02$, $p = .89$) , and a marginally significant two-way interaction between AOT and the evaluation task ($F(2, 92) = 2.77$, $p = .07$). To explore the interaction, I examined the slopes of AOT at each level of argument evaluation tasks. The slope of AOT was significant and positive ($b = 0.18$) in Agreement Evaluation Setting and the slope of AOT was negative ($b = -0.06$) in Half & Half Evaluation Setting, while the slope of AOT was negative ($b = -0.14$) in Disagreement Evaluation Setting. In addition, a spotlight analysis at one standard deviation above the mean of AOT showed a significant difference such

that high AOT respondents had more confidence increase in agreement evaluation setting versus in disagreement and half & half evaluation settings. The results showed that in agreement evaluation setting, people with higher AOT had stronger confidence increase, while in disagreement and half & half evaluation settings, people with higher AOT had stronger confidence decrease. Therefore, people with higher AOT will be better off by viewing arguments from the other side or mixed evaluation settings. Therefore, H7.4 was supported.

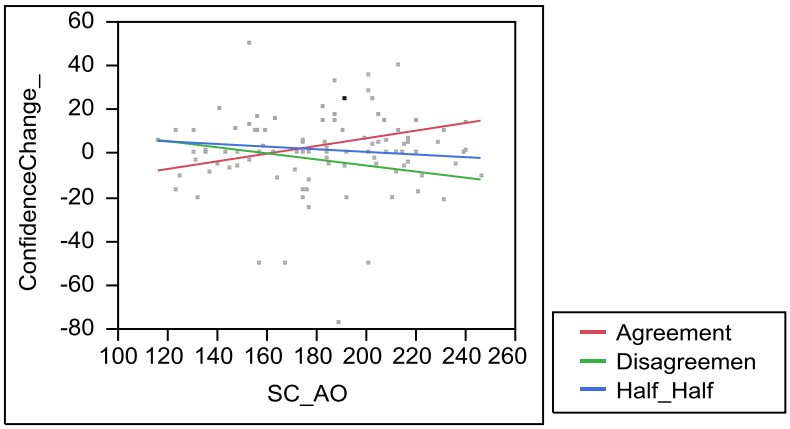

Figure 11. Interaction Effect of Actively Open-minded Thinking and Argument Evaluation Task on Confidence Change

Need for cognition. A regression was performed on confidence change with independent variables (i) NCS, (ii) a categorical variable-argument evaluation task, and (iii) their interaction. The results showed a marginally significant main effect of NCS ($F(1, 93) = 2.82$, $p = .09$) , but no two-way interaction between NCS and argument

evaluation task ($F(2, 92) = 1.12$, $p = .33$). The main effect of NCS ($b = 0.17$) showed that participants with higher need for cognition had less confidence decrease across all three argument evaluation tasks. However, H8.4 was not supported.

Conclusion

Based on the results, I can see that argument generation task and evaluation task do have different impact on attitude polarization. In argument generation task, individual and group think settings lead people to have more extreme positions, while argumentative setting leads people to have more moderate positions. In argument evaluation task, agreement setting leads people to have more extreme positions, while disagreement and half & half settings lead people to have more moderate positions. Therefore, I provide empirical data to show that individual and group think settings do have polarization effect on attitude change and argumentative setting have moderation effect on attitude change. The findings provide some empirical evidence for the literature, especially for the Argumentative Theory (Mercier & Sperber, 2011). However, there are no significant differences in confidence changes among all generation and evaluation settings.

Individual differences did play moderating roles in attitude and confidence changes. There is no moderating effect on attitude polarization in argument generation task. However, the results showed that in agreement evaluation setting, people with higher logical reasoning ability/cognitive reflection test scores have stronger attitude polarization, while in disagreement and half & half evaluation settings, people with

higher logical reasoning ability/cognitive reflection test scores had stronger attitude depolarization/moderation. Therefore, people with higher logical reasoning ability/cognitive reflection test scores will be better off by reading the arguments from the opposite views or mixed arguments by themselves.

Based on the results of confidence change in argument generation task, in group think setting, people with higher cognitive reflection test/need for cognition score had stronger confidence increase; while in argumentative setting, people with higher cognitive reflection test/need for cognition score had stronger confidence decrease. Therefore, people with higher cognitive reflection test/need for cognition score will be better off by arguing with people having different opinions.

Based on the results of confidence change in evaluation task, in agreement evaluation setting, people with higher AOT had stronger confidence increase, while in disagreement and half & half evaluation settings, people with higher AOT had stronger confidence decrease. Therefore, people with higher AOT will be better off by viewing arguments from the other side or mixed evaluation settings.

Overall, viewing otherside arguments and arguing with other people both seem good ways to debias and then avoid attitude polarization and overconfidence. Also, individual differences do moderate the effect of argument generation and evaluation tasks on attitude and confidence change. It seems that, in general, people who are higher in these thinking disposition scales will perform better and depolarize or decrease confidence extremity after viewing arguments from the opposite side or arguing with people holding different views.

Discussion

In the literature, various streams of research try to find out how to improve people's thinking and reasoning skills. They all have their own theories and prediction; however, sometimes the predictions are divergent. In this study, I try to use individual differences to provide some explanations and close the gap. Individual differences can provide some insights about why the results in these studies are not easy to replicate. Not every debiasing strategy can work perfectly on anyone. People in different levels of logical reasoning ability may need to use different debiasing strategy to improve their thinking and make better decisions.

Further, I bring in individual differences into the polarization literature. Try to provide a different angle to interpret how people change their attitude and confidence after argument generation and group discussion. Not everyone in the same group will have the same direction/levels of attitude and confidence changes; even they are going through the same discussion processes. Logical reasoning ability, cognitive reflection test, actively open-minded thinking and need for cognition did moderate this information processes differently. I will replicate these experiments in future studies.

GENERAL DISCUSSION

Researchers studying decision making have largely focused on how the decisions that people make are affected by task characteristics, how labile decisions are with respect to situational factors, and the degree to which preferences are constructed at the time of the decision. However, the fact that many preferences are constructed does not imply that all constructions are equally good. That people differ from each other is obvious. How and why they differ is less clear. Perhaps the most important question of individual differences is whether people are more similar to themselves over time and across situations than they are to others, and whether the variation within a single person across time and situation is less than the variation between people.

If we accept that being "rational" means that "rational" thought helps people achieve their goals (Baron, 2000), then we conclude that, in almost all decision environments, good thinking involves sufficient search for possibilities and evidence, and fairness in the evaluation of the obtained evidence (e.g., Baron, 1985). Even early proponents of constructed preferences conceded that non-situational, individual-level characteristics could exert strong influences on both strategies for approaching a decision problem and on the construction of preferences. Payne, Bettman, and Johnson (1993) note that the expertise of the decision maker could lead to substantial differences in approach and outcome, and that other characteristics of the decision maker could also matter.

An issue that remains open is the degree to which people can develop expertise in the process of decision making itself. Many researchers conclude that expertise is

strongly domain-bound, with little or no far-transfer (see Barnett & Ceci, 2002, for a review). Other researchers, however, argue that certain individual differences that could be classified as forms of "expertise in thinking" have the characteristics of traits, in that they are stable, not domain specific, and can affect entire substrates of thought (e.g., Keith E. Stanovich, 2000).

Interestingly, measures of intelligence, often taken as the benchmark of good thinking, do not assess the avoidance of myside bias and one-sided bias—the foundational skills in the critical thinking literature and in the rational thinking literature aimed at educational applications (Baron, 2000; Kuhn, 1993, 2005; Nussbaum & Sinatra, 2003; Perkins, 1995; Sternberg, 1997, 2001, 2003). Furthermore, I will argue here that there are theoretical reasons to believe that these classes of thinking skills—myside thinking and one-sided thinking—are likely to be unusually dissociated from measures of cognitive ability. Because intelligence is such a ubiquitous associate of cognitive performance, it is noteworthy to identify the types of cognitive tasks that tend to eliminate its predictive power (Stanovich & West, 2008).

In the series of three studies, I found little or no relationship between measures of cognitive ability and bias or argument quality prior to prompting. However, in the first and second studies, I support the view that trait-like dispositions predict fewer biases in reasoning and increased quality of generated arguments, including both-sides evidence, in various constructed preference-like tasks. There is a strong positive relationship between acumen in logical reasoning ability and both argument quality and reduction of common biases in reasoning such as myside bias, especially in argument generation task.

However, in Study 2, I found different results in argument evaluation task. Logical reasoning ability is still the good predictor of argument evaluation ability, but it is also positively related with myside bias in argument evaluation task. The differing findings show that the so-called "confirmation bias" measurements could be indicating different things in argument generation and evaluation tasks. I will need to dig this issue further in future studies.

Furthermore, I found that CRT is also a good predictor of the argument generation quality and the ability to avoid myside bias in argument evaluation task. Also, AOT is a good predictor of argument evaluation ability. Most importantly, cognitive ability is never a good predictor in both argument generation and evaluation tasks, which provide some findings to fill the gap in literature. However, these studies showed that some thinking dispositions are good predictors of thinking ability and debiasing ability.

For standard theories of reasoning, the confirmation bias is no more than a flaw of reasoning, however, for the argumentative theory, the confirmation bias is a consequence of the function of reasoning and hence a feature of reasoning when used for the argument production (Mercier & Sperber, 2011). The argumentative theory puts such well-known demonstrations of "irrationality" in a novel perspective. They propose that human reasoning is a remarkably efficient specialized device adapted to a certain type of social and cognitive interaction at which it excels, instead of a profoundly flawed general mechanism.

According to the Argumentative Theory (Mercier & Sperber, 2011), when one is alone (individual) or with people who hold similar views (group think), one's arguments

will not be critically evaluated or challenged. This is when the confirmation bias is most likely to lead to poor outcomes and suboptimal decisions. However, when reasoning is used in a more argumentative context—that is, in arguments among people who disagree with each other but have a common interest in the truth—the confirmation bias contributes to an efficient form of division of cognitive labor. In that situation, a group can work more efficiently to solve a problem if each individual looks mostly for arguments supporting a given solution, and brings these arguments to be tested by other group members.

The argumentative theory also proposes that the "confirmation bias" should be active only in the argument generation but not in argument evaluation tasks, no matter in discussions or group performance. They argue that people tend to be more objective in argument evaluation than in argument generation (Mercier & Sperber, 2011). They believe that the main function of reasoning is argumentative, and that's the reason why the past research in reasoning can't find good outcomes from abstract reasoning tasks. Therefore, people improve their argument quality significantly when they have to debate with other people (Blanchette & Dunbar, 2001; Kuhn, Shaw, & Felton, 1997; Mercier & Sperber, 2011).

Biased search, interpretation and memory contribute to overconfidence, attitude polarization in the face of the contrary evidence, belief perseverance, the irrational primacy effect and illusory correlation. Further, poor thinking could be caused by too little search, by overconfidence, and by biases in favor of possibilities that are favored initially. On the contrary, good thinking consists of (1) thorough search based on the importance of the question, (2) appropriate confidence in the quality of thinking done,

and (3) fairness to all the possibilities. As being said, pushing people away from overconfidence and extreme opinions will lead to the path of good thinking.

Based on the results of Study 3, we can see that argument generation task and evaluation task do have different impacts on attitude polarization. In argument generation task, individual and group think settings lead people to have more extreme positions, while argumentative setting leads people to have more moderate positions. In argument evaluation task, agreement setting leads people to have more extreme positions, while disagreement and half & half settings lead people to have more moderate positions. Therefore, I provide empirical data to show that individual and group think settings do have polarization effect on attitude change and argumentative setting have moderation effect on attitude change. Our findings provide some empirical evidence for the literature, especially for the argumentative theory (Mercier & Sperber, 2011). More interestingly, the findings show that argumentation settings do decrease attitude extremity not only in argument generation task, but also in argument evaluation task, which is different from the predictions of Mercier and Sperber (2011).

Most importantly, individual differences do play moderating roles in attitude and confidence changes. There is no moderating effect on attitude polarization in argument generation task. However, the results showed that in agreement evaluation setting, people with higher logical reasoning ability/cognitive reflection test scores have stronger attitude polarization, while in disagreement and half & half evaluation settings, people with higher logical reasoning ability/cognitive reflection test scores had stronger attitude depolarization/moderation. Therefore, people with higher logical reasoning

ability/cognitive reflection test scores will be better off by reading the arguments from the opposite views by themselves.

Based on the results of confidence change in generation task, in group think setting, people with higher cognitive reflection test/need for cognition score had stronger confidence increase; while in argumentative setting, people with higher cognitive reflection test/need for cognition score had stronger confidence decrease. Therefore, people with higher cognitive reflection test/need for cognition score will be better off by arguing with people having different opinions.

Based on the results of confidence change in evaluation task, in agreement evaluation setting, people with higher AOT had stronger confidence increase, while in disagreement and half & half evaluation settings, people with higher AOT had stronger confidence decrease. Therefore, people with higher AOT will be better off by viewing arguments from the other side or mixed evaluation settings.

Overall, viewing other-side arguments and arguing with other people both seem good ways to debias and then decrease attitude and confidence extremity. Most interestingly and importantly, individual differences do interact with the effect of the generation and evaluation tasks on attitude and confidence change. The results showed that individuals do perform differently; therefore, I provided some critical evidence for the literature. In the series of three studies, the results show that certain individual differences could be classified as forms of "expertise in thinking" have the characteristics of traits, in that they are stable, not domain specific, and can affect entire substrates of thought (e.g., Keith E. Stanovich, 2000).

REFERENCES

Ackerman, P. L. (1994). Intelligence, attention, and learning: Maximal and typical performance. In D. K. Detterman (Ed.), *Current topics in human intelligence* (pp. 1-27). Norwood, NJ: Ablex.

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence, 22*(2), 227-257.

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*(2), 219-245.

Ackerman, P. L., & Kanfer, R. (2004). Cognitive, affective, and conative aspects of adult intellect within a typical and maximal performance framework. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrated perspectives on intellectual functioning* (pp. 119-141). Mahwah, NJ: Erlbaum.

Ackerman, P. L., Kyllonen, P. C., & Roberts, R. D. (Eds.). (1999). *Learning and individual differenes: Process, trait, and content determinants*. Washington, DC: American Psychological Association.

Adler, J. E. (1984). Abstraction is uncooperative. *Journal for the Theory of Social Behaviour, 14*(2), 165-181.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R. (1991). Is human cognition adaptive?. *Behavioral and Brain Sciences, 14*(03), 471-485.

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin, 110*(3), 486-498.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612-637.

Baron, J. (1985). *Rationality and Intelligence*. New York: Cambridge University Press.

Baron, J. (1988). *Thinking and Deciding*. Cambridge, England: Cambridge University Press.

Baron, J. (1991). Beliefs about thinking. In J. F. Voss, D. N. Perkins & J. W. Segal (Eds.), *Informal reasoning and education*. Hillsdale, NJ: Erlbaum.

Baron, J. (1993). Why Teach Thinking?-An Essay. *Applied Psychology, 42*(3), 191-214.

Baron, J. (1994). *Thinking and deciding (2nd ed.)*. New York: Cambridge University Press.

Baron, J. (1995). Myside bias in thinking about abortion. *Thinking and Reasoning, 1*, 221-235.

Baron, J. (2000). *Thinking and deciding*. New York: Cambridge University Press.

Bell, D. E., Raiffa, H., & Tversky, A. (Eds.). (1988). *Decision making: Descriptive, normative and prescriptive interactions*. New York, NY: Cambridge University Press.

Blanchette, I., & Dunbar, K. (2001). Analogy use in naturalistic settings: The influence of audience, emotion, and goals. *Memory & Cognition, 29*(5), 730-735.

Bornstein, B. H. (2004). The impact of different types of expert scientific testimony on mock jurors' liability verdicts. *Psychology, Crime & Law, 10*(4), 429-446.

Braman, E. (2009). *Law, politics, and perception: How policy preferences influence legal reasoning*: University of Virginia Press.

Bratman, M. E., Israel, D. J., & Pollack, M. E. (1991). Plans and resource-bounded practical reasoning. In J. Cummins & J. Pollock (Eds.), *Philosophy and AI: Essays at the interface* (pp. 7-22). Cambridge, MA: MIT Press.

Brody, N. (1992). *Intelligence (2nd ed.)*. San Diego, CA: Academic Press.

Brody, N. (1997). Intelligence, schooling, and society. *American Psychologist, 52*(10), 1046-1050.

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology, 92*(5), 938-956.

Buss, D. M. (1991). Evolutionary personality psychology. *Annual Review of Psychology, 42*(1), 459.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*(1), 116-131.

Cacioppo, J. T., & Petty, R. E. (1984). The need for cognition: Relationship to attitudinal processes. In R. P. McGlynn, J. E. Maddux, C. D. Stoltenberg & J. H. Harvey (Eds.), *Social perception in clinical and counseling psychology* (pp. 91-119): Lubbock: Texas Tech University.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197-253.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306-307.

Cacioppo, J. T., Petty, R. E., & Morris, K. J. (1983). Effects of need for cognition on message evaluation, recall, and persuasion. *Journal of Personality and Social Psychology, 45*(4), 805-818.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review, 97*(3), 404-431.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, England: Cambridge University Press.

Carroll, J. B. (1995). The higher-stratum structure of cognitive abilities: Current evidence supports g and about ten broad factors. In H. Nyborg (Ed.), *The scientific study of general intelligence: tribute to Arthur Jensen*. Oxford, England: Elsevier Science/Pergamon Press.

Carroll, J. B. (1995). On methodology in the study of cognitive abilities. *Multivariate Behavioral Research, 30*(3), 429-452.

Carroll, J. B. (1997). Psychometrics, intelligence, and public perception. *Intelligence, 24*, 25-52.

Carroll, J. B. (2003). The higher-stratum structure of cognitive abilities: Current evidence supports g and about ten broad factors. In H. Nyborg (Ed.), *The scientific study of general intelligence: tribute to Arthur Jensen*. Oxford, England: Elsevier Science/Pergamon Press.

Caryl, P. G. (1994). Early event-related potentials correlate with inspection time and intelligence. *Intelligence, 18*, 15-46.

Cherniak, C. (1986). *Minimal rationality*. Cambridge, MA: MIT Press.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences, 4*(03), 317-331.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making, 4*, 20-33.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence, 32*(3), 277-296.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7*(12), 547-552.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58*(1), 1-73.

Costa, P. T., & McCrae, R. R. (1992). *Neo Pi-R: Professional manual/with bibliography*. Odessa, FL: Psychological Assessment Resources.

Coutinho, S. A., & Woolery, L. M. (2004). The need for cognition and life satisfaction among college students. *College Student Journal, 38*(2), 203-206.

Cronbach, L. J. (1949). *Essentials of psychological testing*. Oxford, England: Harper.

Davis, F. D., Lohse, G. L., & Kottemann, J. E. (1994). Harmful effects of seemingly helpful information on forecasts of stock earnings. *Journal of Economic Psychology, 15*(2), 253-267.

Dawes, R. M. (1976). Shallow psychology. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior* (pp. 3-11). Hillsdale: Erlbaum.

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition, 106*(3), 1248-1299.

Deary, I. J. (1995). Auditory inspection time and intelligence: What is the direction of causation? *Developmental Psychology, 31*(2), 237-250.

Deary, I. J. (2001). *Intelligence: A very short introduction*. Oxford, UK: Oxford University Press.

Deary, I. J., & Stough, C. (1996). Intelligence and inspection time: Achievements, prospects, and problems. *American Psychologist, 51*(6), 599-608.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish mental surveys of 1932 and 1947. *Journal of Personality and Social Psychology, 86*(1), 130-147.

Dennett, D. C. (1978). *Brainstorms: Philosophical essays on mind and psychology*. Cambridge: Ma: MIT Press.

Dennett, D. C. (1987). *The intentational stance*. Cambridge, MA: MIT Press.

DeShon, R. P., Smith, M. R., Chan, D., & Schmitt, N. (1998). Can racial differences in cognitive test performance be reduced by presenting problems in a social context? *Journal of Applied Psychology, 83*(3), 438-451.

Detterman, D. K. (1994). Intelligence and the brain. In P. A. Vernon (Ed.), *The neuropsychology of individual differences* (pp. 35-57). San Diego: Academic Press.

Dewey, J. (1922). *Human nature and conduct: An introduction to social psychology*. New York, NY: H. Holt and Company.

Dole, J. A. S. G. M. (1998). Reconceptalizing change in the cognitive construction of knowledge. *Educational Psychologist, 33*(2/3), 109.

Dougherty, T. M., & Haith, M. M. (1997). Infant expectations and reaction time as predictors of childhood speed of processing and IQ. *Developmental Psychology, 33*(1), 146-155.

Eagly, A. H., Kulesa, P., Brannon, L. A., Shaw, K., & Hutson-Comeaux, S. (2000). Why counterattitudinal messages are as memorable as proattitudinal messages: The importance of active defense against attack. *Personality and Social Psychology Bulletin, 26*(11), 1392-1408.

Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology, 71*(1), 5-24.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19-23.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309-331.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9-26). New York: W. H. Freeman.

Epstein, S., & Meier, P. (1989). Constructive thinking: A broad coping variable with specific components. *Journal of Personality and Social Psychology, 57*(2), 332-350.

Evans, J. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin, 128*(6), 978-996.

Evans, J. S. B. T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology, 75*(4), 451.

Evans, J. S. B. T. (1989). *Bias in human reasoning: Causes and consequences.* Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Evans, J. S. B. T. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning, 13*(4), 321-339.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*(1), 255-278.

Evans, J. S. B. T. (2010). *Thinking twice: Two minds in one brain*. Oxford, Oxford University Press.

Evans, J. S. B. T., Barston, J., & Pollard, P. (1983). On the cognition between logic and belief in syllogistic reasoning. *Memory & Cognition, 11*(3), 295-306.

Evans, J. S. B. T., & Frankish, K. (Eds.). (2009). *In two minds: Dual processes and beyond*. Oxford, England: Oxford University Press.

Evans, J. S. B. T., & Over, D. E. (1996). Rationality in the selection task: Epistemic utility versus uncertainty reduction. *Psychological Review, 103*(2), 356-363.

Facione, P. (2011). *Think Critically*. Upper Saddle River, NJ: Pearson.

Facione, P. A., Sánchez, C. A., Facione, N. C., & Gainen, J. (1995). The disposition toward critical thinking. *The Journal of General Education, 44*(1), 1-25.

Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review, 90*(3), 239-260.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1979). Subjective sensitivity analysis. *Organizational Behavior and Human Performance, 23*(3), 339-359.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives, 19*(4), 25-42.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science, 15*(6), 373-378.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science, 7*(4), 237-241.

Galotti, K. M. (1989). Approaches to studying formal and everyday reasoning. *Psychological Bulletin, 105*(3), 331-351.

Gauthier, K. J., Christopher, A. N., Walter, M. I., Mourad, R., & Marek, P. (2006). Religiosity, religious doubt, and the need for cognition: Their interactive relationship with life satisfaction. *Journal of Happiness Studies, 7*(2), 139-154.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review, 103*(3), 592-596.

Gilovich, T. (1993). *How we know what isn't so: the fallibility of human reason in everyday life*. New York, NY: The Free Press.

Greenhoot, A. F., Semb, G., Colombo, J., & Schreiber, T. (2004). Prior beliefs and methodological concepts in scientific reasoning. *Applied Cognitive Psychology, 18*(2), 203-221.

Greenwald, A. G. (1969). The open-mindedness of the counterattitudinal role player. *Journal of Experimental Social Psychology, 5*(4), 375-388.

Guillermo, C., & Martín, L. (2010). Correlations of cognitive reflection with judgments and choices. *Judgment and Decision Making, 5*(3), 182.

Handley, S. J., Capon, A., Beveriddge, M., Dennis, I., & Evans, J. S. B. T. (2004). Working memory, inhibitory control and the development of children's reasoning. *Thinking & Reasoning, 10*(2), 175-195.

Harman, G. (1995). Rationality. In E. E. Smith & D. N. Osherson (Eds.), *Thinking: An invitation to cognitive science (2nd ed., Vol. 3)*. Cambridge, MA: MIT Press.

Hastie, R., & Park, B. (2005). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. In D. L. Hamilton (Ed.), *Social Cognition: key readings*. New York: Psychology Press.

Henderson, W. D. (2004). The LSAT, Law schoo exams and meritocracy: The surprising and understanding role of test-taking speed. *Faculty Publications*, Paper 344.

Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin, 118*(2), 248-271.

Hindery, R. (2001). *Indoctrination and self-deception or free and critical thought?* Lewiston, NY: Edwin Mellen Press.

Honderich, T. (1995). *The Oxford companion to philosophy*. New York, NY: Oxford University Press.

Horgan, T., & Tienson, J. (1993). Levels of description in nonclassical cognitive science. In C. Hookway & D. Peterson (Eds.), *Philosophy and cognitive science* (pp. 159-188). Cambridge, England: Cambridge University Press.

Hunt, E. (1987). The next word on verbal ability. In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 347-392). Norwood, NJ: Ablex.

Hunt, E. (1999). Intelligence and human resources: Past, present, and future. In P. L. Ackerman, P. C. Kyllonen & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 3-30). Washington, DC: American Psychology Association.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Greenwood Publishing.

Kahneman, D. (2000). A psychological point of view: Violations of rational rules as a diagnostic of mental processes. *Behavioral and Brain Sciences, 23*(05), 681-683.

Kahneman, D., & Frederick, S. (2002). Representative revisited: Attribute substitution in intuitive judgment. In D. G. T. Gilovich & D. Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (pp. 49-81). New York: Cambridge University Press.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*(3), 430-454.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103*(3), 582-591.

Kane, M., & Engle, R. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9*(4), 637-671.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*(2), 189-217.

Keating, D. P. (1990). Charting pathways to the development of expertise. *Educational Psychologist, 25*(3/4), 243.

Kitcher, P. (1993). *The advancement of science: Science without legend, objectivity without illusions*. New York: Oxford University Press.

Klaczynski, P. A. (1997). Bias in adolescents' everyday reasoning and its relationship with intellectual ability, personal theories, and self-serving motivation. *Developmental Psychology, 33*(2), 273-283.

Klaczynski, P. A., & Gordon, D. H. (1996). Self-serving influences on adolescents' evaluations of belief-relevant evidence. *Journal of Experimental Child Psychology, 62*(3), 317-339.

Klaczynski, P. A., Gordon, D. H., & Fauth, J. (1997). Goal-oriented critical reasoning and individual differences in critical reasoning biases. *Journal of Educational Psychology, 89*(3), 470-485.

Klaczynski, P. A., & Lavallee, K. L. (2005). Domain-specific identity, epistemic regulation, and intellectual ability as predictors of belief-biased reasoning: A dual-process perspective. *Journal of Experimental Child Psychology, 92*(1), 1-24.

Klaczynski, P. A., & Narasimham, G. (1998). Development of scientific reasoning biases: Cognitive versus ego-protective explanations. *Developmental Psychology, 34*(1), 175-187.

Klaczynski, P. A., & Robinson, B. (2000). Personal theories, intellectual ability, and epistemological beliefs: Adult age differences in everyday reasoning biases. *Psychology and Aging, 15*(3), 400-416.

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review, 107*(4), 852-884.

Koehler, D., & James, G. (2010). Probability matching and strategy availability. *Memory & Cognition, 38*(6), 667.

Kogan, N., & Wallach, M. A. (1966). Modification of a judgmental style through group interaction. *Journal of Personality and Social Psychology, 4*(2), 165-174.

Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology, 83*(1), 26-52.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 107-118.

Kruglanski, A. W. (1990). Lay epistemic theory in social-cognitive psychology. *Psychological Inquiry, 1*(3), 181-197.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: 'Seizing' and 'freezing.'. *Psychological Review, 103*(2), 263-283.

Kuhn, D. (1991). *The skills of argument*. New York, NY: Cambridge University Press.

Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review, 62*(2), 155-179.

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education, 77*(3), 319-337.

Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher, 28*(2), 16-46.

Kuhn, D. (2001). How do people know? *Psychological Science, 12*(1), 1-8.

Kuhn, D. (2005). *Education for thinking*. Cambridge, MA: Harvard University Press.

Kuhn, D., Shaw, V. F., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Conition and Instruction, 15*(3), 287-315.

Larrick, R. P. (2004). Debiasing In D. J. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making* (pp. 316-337). Malden, MA: Blackwell Publishing.

LeBoeuf, R. A., & Shafir, E. (2003). Deep thoughts and shallow frames: on the susceptibility to framing effects. *Journal of Behavioral Decision Making, 16*(2), 77-92.

Levelt, W. (1995). Chapters of psychology. In R. L. Solso & D. W. Massaro (Eds.), *The science of the mind: 2001 and beyond* (pp. 184-202). New York: Oxford University Press.

Levi, I. (1983). Who commits the base rate fallacy? *Behavioral and Brain Sciences, 6*(03), 502-506.

Levinson, S. C. (1995). Interactional biases in human thinking. In E. Goody (Ed.), *Social intelligence and interaction* (pp. 221-260). Cambridge, England: Cambridge University Press.

Lichtenstein, S., & Slovic, P. (2006). *The construction of preference*. New York: Cambridge University Press.

Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology, 58*(1), 259-289.

Lieberman, M. D. (2009). What zombies can't do: A social cognitive neuroscience approach to the irreducibility of reflective consciousness. In J. S. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 293-316). Oxford, England: Oxford University Press.

Lohman, D. F. (2000). Complex information processing and intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 285-340). New York, NY: Cambridge University Press.

Lopes, L. L. (1981). Performing competently. *Behavioral and Brain Sciences, 4*(03), 343-344.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*(11), 2098-2109.

Lubinski, D. (2000). Scientific and social significance of assessing individual differences: `Sinking shafts at a few. *Annual Review of Psychology, 51*(1), 405.

Lubinski, D., & Humphreys, L. G. (1997). Incorporating general intelligence into epidemiology and the social sciences. *Intelligence, 24*(1), 159-201.

Macdonald, R. (1986). Credible conceptions and implausible probabilities. *British Journal of Mathematical and Statistical Psychology, 39*, 15-27.

Mackintosh, N. J. (1998). *IQ and human intelligence*. New York, NY: Oxford University Press.

Margolis, H. (1987). *Patterns, thinking, and cognition*. Chicago: University of Chicago Press.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of the adult intelligence. (5th ed.)*. Oxford, England: Williams & Wilkins.

Matthews, G., Davies, D. R., Westerman, S. J., & Stammers, R. B. (2000). *Human performance: Cognition, stress and individual differences*. London: Psychology Press.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional intelligence: Science and myth*. Boston: MIT Press.

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*(02), 57-74.

Messick, S. (1984). The nature of cognitive styles: Problems and promise in educational practice. *Educational Psychologist, 19*(2), 59-74.

Messick, S. (1994). The matter of style: Manifestations of personality in cognition, learning, and teaching. *Educational Psychologist, 29*(3), 121.

Moshman, D. (1994). Reason, reasons and reasoning: A constructivist account of human rationality. *Theory & Psychology, 4*(2), 245-260.

Newell, A. (1982). The knowledge level. *Artificial Intelligence, 18*(1), 87-127.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newstead, S. E., Handley, S. J., Harley, C., Wright, H., & Farrelly, D. (2004). Individual differences in deductive reasoning. *Quarterly Journal of Experimental Psychology: Section A, 57*(1), 33-60.

Nickerson, R. S. (1987). *Why teach thinking?* New York, NY: W H Freeman/Times Books/Henry Holt & Co.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175-220.

Nickerson, R. S. (2004). Teaching reasoning. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning*. Cambridge, UK: Cambridge University Press.

Nisbett, R. E. (1993). Violence and U.S. regional culture. *American Psychologist, 48*(4), 441-449.

Norris, S. P. (1992). *Ceneralizability of critical thinking: Multiple perspectives on an educational ideal*. New York: Teachers College Press.

Norris, S. P., & Ennis, R. H. (1989). *Evaluating critical thinking. The practitioners' guide to teaching thinking series*. Pacific Grove, CA: Critical Thinking Press and Software.

Nussbaum, E. M., & Sinatra, G. M. (2003). Argument and conceptual engagement. *Contemporary Educational Psychology, 28*(3), 384-395.

Oaksford, M., & Chater, N. (1993). Reasoning theories and bounded rationality. In K. Manktelow & D. E. Over (Eds.), *Rationality: Psychological and philosophical perspectives*. London: Routledge.

Oaksford, M., & Chater, N. (1995). Theories of reasoning and the computational explanation of everyday inference. *Thinking & Reasoning, 1*, 121-152.

Obrecht, N., Chapman, G., & Gelman, R. (2009). An encounter frequency account of how experience affects likelihood estimation. *Memory & Cognition, 37*(5), 632.

Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization, 72*(1), 147-152.

Osherson, D. N. (1995). Probability judgment. In E. E. Smith & D. N. Osherson (Eds.), *An invitation to cognitive science-2nd Edition, Vol. 3: Thinking*. Cambridge, MA: The MIT Press.

Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology, 29*(3), 261-265.

Paese, P. W., & Sniezek, J. A. (1991). Influences on the appropriateness of confidence in judgment: Practice, effort, information, and decision-making. *Organizational Behavior and Human Decision Processes, 48*(1), 100-130.

Parker, A. M., & Fischhoff, B. (2005). Decision-making competence: External validation through an individual-differences approach. *Journal of Behavioral Decision Making, 18*(1), 1-27.

Paul, R. W. (1984). Critical thinking: Fundamental to education for a free society. *Educational Leadership, 42*(1), 4.

Paul, R. W. (1987). *Dialogical thinking: Critical thought essential to the acquisition of rational knowledge and passions*. New York, NY: W H Freeman/Times Books/Henry Holt & Co.

Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology, 60*(2), 307-317.

Payne, J. W., Bettman, J. R., & Johnson, E. (1993). *The adaptive decision maker*. New York: Cambridge Universtiy Press.

Perkins, D. N. (1985). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology, 77*(5), 562-571.

Perkins, D. N. (1995). *Outsmarting IQ: The emerging science of learnable intelligence*. New York: Free Press.

Perkins, D. N., Bushey, B., & Faraday, M. (1986). *Leraning to reason*. Unpublished manuscript, Cambridge, MA: Harvard Graduate School of Education.

Perkins, D. N., Faraday, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. F. Voss, D. N. Perkins & J. W. Segal (Eds.), *Informal reasoning and education*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Perkins, D. N., Jay, E., & Tishman, S. (1993). Beyond abilities: A dispositional theory of thinking. *Merrill-Palmer Quarterly, 39*(1), 1-21.

Perkins, D. N., & Ritchhart, R. (2004). When is good thinking? In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, Emotion, and Cognition: Integrative perspectives on intellectual functioning and development*. Mahwah, NJ: Lawrence Erlbaum Associates.

Peterson, D. K., & Pitz, G. F. (1988). Confidence, uncertainty, and the use of information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(1), 85-92.

Plous, S. (1993). *The psychology of judgment and decision making*. New York, NY, England: Mcgraw-Hill Book Company.

Pollock, J. L. (1995). *Cognitive carpentry: A blueprint for how to build a person*. Cambridge, MA: MIT Press.

Polya, G. (1945). *How to solve it: A new aspect of mathematical method*. Princeton, NJ: Princeton University Press.

Pomerantz, E. M., Chaiken, S., & Tordesillas, R. S. (1995). Attitude strength and resistance processes. *Journal of Personality and Social Psychology, 69*(3), 408-419.

Pylyshyn, Z. (1984). *Computation and cognition*. Cambridge, MA: MIT Press.

Redlawsk, D. P. (2002). Hot cognition or cool consideration? Testing the effects of motivated reasoning on political decision making. *The Journal of Politics, 64*(04), 1021-1044.

Ritchart, R., & Perkins, D. N. (2005). Learning to think: The challenges of teaching thinking. In H. K. & R. Morrison (Eds.), *The cambridge handbook of thinking and reasoning*. New York: Cambridge University Press.

Ritchhart, R. (2002). *Intellectual character: What it is, why it matters, and how to get it*. San Francisco, CA: John Wiley & Sons, Inc.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). Criteria for scale selection and evaluation. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 1-15). San Diego, CA: Academic Press.

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: an illusion of explanatory depth. *Cognitive Science, 26*(5), 521-562.

Sá, W. C., Kelley, C. N., Ho, C., & Stanovich, K. E. (2005). Thinking about personal theories: individual differences in the coordination of theory and evidence. *Personality and Individual Differences, 38*(5), 1149-1161.

Sá, W. C., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 91*(3), 497-510.

Sadowski, C. J., & Gülgös, S. (1996). Elaborative processing mediates the relationship between need for cognition and academic performance. *The Journal of Psychology, 130*(3), 303-307.

Sadowski, C. J., & Gulgoz, S. (1992). Internal consistency and test-retest reliability of the need for cognition scale. *Perceptual and Motor Skills, 74*(2), 610-610.

Scheffler, I. (1991). *In praise of the cognitive emotions and other essays in the philosophy of education*. New York, NY: Routledge.

Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. *Educational Psychologist, 39*(1), 19-29.

Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. Mahweh, NJ: Lawrence Erlbaum Associates.

Shafir, E., & Tversky, A. (1995). Decision making. In E. E. Smith & D. N. Osherson (Eds.), *Thinking: An invitation to cognitive science, Vol.3 (2nd ed.). An invitation to cognitive science (2nd ed.)* (pp. 77-100). Cambridge, MA, US: The MIT Press.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*(1), 99-118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*(2), 129-138.

Simon, H. A. (1957). *Models of man*. New York: Wiley.

Sinatra, G. M., & Pintrich, P. R. (2003). The role of intentions in conceptual change learning. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change* (pp. 1-18). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Sloman, A. (1993). The mind as a control system. In H. C. & D. Peterson (Eds.), *Philosophy and cognitive science* (pp. 69-110). Cambridge, England: Cambridge University Press.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*(1), 3-22.

Sloman, S. A. (2002). Two systems of reasoning. In D. G. Gilovich & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology, 15*(2), 201-292.

Stankov, L., & Dunn, S. (1993). Physical substrata of mental energy: Brain capacity and efficiency of cerebral metabolism. *Learning and Individual Differences, 5*(3), 241-257.

Stanovich, K. E. (2000). *Who is rational?* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Stanovich, K. E. (2002). Rationality, intelligence, and levels of analysis in cognitive science: Is dysrationalia possible? In R. J. Sternberg (Ed.), *Why smart people can be so stupid?* (pp. 124-158). New Haven, CT: Yale University Press.

Stanovich, K. E. (2003). The fundamental computational biases of human cognition: Heuristics that (sometimes) impair decision making and problem solving. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 291-342). New York: Cambridge University Press.

Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago: IL: University of Chicago Press.

Stanovich, K. E. (2009a). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. S. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond*. Oxford, England: Oxford University Press.

Stanovich, K. E. (2009b). *What intelligence tests miss: The psychology of rational thought*. New Haven: Yale University Press.

Stanovich, K. E. (2011). *Rationality and the reflective mind*. New York: Oxford University Press.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology, 89*(2), 342-357.

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127*(2), 161-188.

Stanovich, K. E., & West, R. F. (1999). Discrepancies between normative and descriptive models of decision making and the understanding/acceptance principle. *Cognitive Psychology, 38*(3), 349-385.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(05), 645-665.

Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning, 13*(3), 225-247.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94*(4), 672-695.

Stein, E. (1996). *Without good reason: The rationality debate in philosophy and cognitive psychology*. Oxford, U.K.: Oxford University Press.

Sterelny, K. (1990). *The representational theory of mind: An introduction*. Oxford, England: Basil Blackwell.

Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology, 49*(3), 607-627.

Sternberg, R. J. (1989). *The triarchic mind: A new theory of human intelligence*. New York: Penguin Books.

Sternberg, R. J. (1997a). The concept of intelligence and its role in lifelong learning and success. *American Psychologist, 52*(10), 1030-1037.

Sternberg, R. J. (1997b). Inspection time for inspection time: A reply to Deary and Stough. *American Psychologist, 52*(10), 1144-1147.

Sternberg, R. J. (2000). The concept of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence*. New York, NY: Cambridge University Press.

Sternberg, R. J. (2001). Why shools should teach for wisdom: The balance theory of wisdom in educational settings. *Educational Psychologist, 36*(4), 227-245.

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. Cambridge, UK: Cambridge University Press.

Sternberg, R. J. (1988). *The nature of creativity: Contemporary psychological perspectives*. New York, NY: Cambridge University Press.

Stich, S. P. (1990). *The fragmentation of reason*. Cambridge, MA: MIT Press.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science, 50*(3), 755-769.

Taylor, S. E. (1981). The interface of cognitive and social psychology. In J. H. Harvey (Ed.), *Cognition, Social behavior, and the environment* (pp. 189-211). Hillsdale: Erlbaum.

Toplak, M., West, R., & Stanovich, K. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39*(7), 1275-1289.

Toplak, M. E., & Stanovich, K. E. (2002). The domain specificity and generality of disjunctive reasoning: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 94*(1), 197-209.

Troldahl, V. C., & Powell, F. A. (1965). A short-form dogmatism scale for use in field studies. *Social Forces, 44*(2), 211-214.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Unsworth, N., & Engle, R. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory & Cognition, 33*(2), 213-220.

Vernon, P. A. (1990). The use of biological measures to estimate behavioral intelligence. *Educational Psychologist, 25*(3/4), 293.

Vernon, P. A. (1993). *Biological approaches to the study of human intelligence*. Norwood, NJ: Ablex.

Vinokur, A., & Burnstein, E. (1978). Depolarization of attitudes in groups. *Journal of Personality and Social Psychology, 36*(8), 872-885.

Voss, J. F., Perkins, D. N., & Segal, J. W. (1991). *Informal reasoning and education*. Hillsdale, NJ: England: Lawrence Erlbaum Associates, Inc.

Wade, C., & Tavris, C. (1993). *Critical and creative thinking: The case of love and war*. New York: Harper Collins.

West, R. F., & Stanovich, K. E. (1991). The incidental acquisition of information from reading. *Psychological Science, 2*(5), 325-330.

Willingham, D. T. (2008). Critical thinking: Why is it so hard to teach? *Arts Education Policy Review, 109*(4), 21-32.

Zechmeister, E. B., & Johnson, J. E. (1992). *Critical thinking: a functional approach*. Belmont, CA: Thomson Brooks/Cole Publishing Co.

APPENDIX A

CONSENT FORM

Protocol Title: Learning, Expertise, and Decision Quality
Principal Investigator: Eric Eisenstein, PhD
Fox School of Business, Marketing Department
Telephone: 215-204-7039
Temple IRB protocol #: 11982

You are being asked to participate in a research study conducted by Eric Eisenstein, Ph.D., a faculty member at the Fox School of Business, Temple University, along with Shih-Ching Wang, a Ph.D. student. You are eligible to participate in this study if you are at least 18 years of age.

The purpose of this study is to gain a better understanding of how people make decisions and process information in a complex environment. If you agree to participate in this study, we will ask you read a series of questions and to answer them. Most questions pertain to how you make decisions in a variety of contexts, but some questions are background questions that provide insight into your habits, personality, and thinking styles. Participating indicates that you grant us permission to check or corroborate your answers against university records. The entire study will require less than 0.5 hour to complete.

Risks and Benefits: No risks are anticipated beyond those encountered in day-to-day life. However, if the study causes you to experience distress or discomfort, you may contact the Temple Community Counseling Center at 215-204-1591. You may derive a direct benefit from participating in that you will have an opportunity to practice a set of skills that are important to make better decisions. From a societal point of view, this research is important because it may lead to the development of better ways to help people make sound decisions.

Compensation: You will receive 2.5 dollars via MTurk if you complete the survey thoroughly. (Notice: Every question is required to answer and contains minimum time spending on each page, which means the next button won't show up before the minimum time spending passes. So please answer each question seriously and NO RUSH!)

Taking part is voluntary. Your decision whether or not to participate will not affect your current or future relations with Temple University. If you decide to participate, you are free to withdraw at any time without affecting those relationships.

Confidentiality: Although the study team has placed safeguards to maintain the confidentiality of my personal information, there is always a potential risk of an unpermitted disclosure. To that degree, all documents and information pertaining to this research study will be kept confidential, unless required by applicable federal, state, and local laws and regulations to be disclosed. The researchers may contact you in the future to participate in additional research in this stream. I understand the records and data generated by the study may be reviewed by Temple University and its agents, the study sponsor or the sponsor's agents (if applicable), and/or governmental agencies to assure proper conduct of the study and compliance with regulations. I understand that the results of this study may be published. If any data is published, I will not be identified by name.

Questions: Please ask any questions you have by emailing Dr. Eisenstein at studyquestion@temple.edu. If you have any questions about your rights as a research subject, you may contact the Institutional Review Board Coordinator at (215) 707-3390. The IRB Coordinator may also be reached by email: IRB@Temple.edu or regular mail: Institutional Review Board Coordinator Temple University Research Administration Student Faculty Conference Center 3340 North Broad Street-Suite 304 Philadelphia, PA 19140

Statement of Consent: I have read this consent form, and I have received answers to all of my questions. Clicking the "Continue" link below serves as my signature, which indicates that I understand this document and that I freely consent to participate in this study. By providing responses to the survey, I understand that this is an agreement to allow the researcher to use my responses as described above.

This information is approved by Temple University for public display and is associated with project 11982

APPENDIX B

NEED FOR COGNITION SCALE

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)
4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities. (reverse)
5. I try to anticipate and avoid situations where there is likely chance I will have to think in depth about something. (reverse)
6. I find satisfaction in deliberating hard and for long hours.
7. I only think as hard as I have to. (reverse)
8. I prefer to think about small, daily projects to long-term ones. (reverse)
9. I like tasks that require little thought once I've learned them. (reverse)
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)
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)
17. It's enough for me that something gets the job done; I don't care how or why it works. (reverse)
18. I usually end up deliberating about issues even when they do not affect me personally.

APPENDIX C

ACTIVE OPEN-MINDED THINKING SCALE

Flexible Thinking Subscale

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

Openness-Values Subscale

11. I believe that laws and social policies should change to reflect the needs of a changing world.
12. I believe letting students hear controversial speakers can only confuse and mislead them. (reverse)
13. I believe we should look to our religious authorities for decisions on moral issues. (reverse)
14. I consider myself broad-minded and tolerant of other people's lifestyles.
15. I believe that loyalty to one's ideals and principles is more important than "open-mindedness." (reverse)
16. I believe that the "new morality" of permissiveness is no morality at all. (reverse)
17. 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)
18. I believe that the different ideas of right and wrong that people in other societies have may be valid for them.

Dogmatism Subscale

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

Categorical Thinking Subscale

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

Belief Identification Subscale

31. What beliefs you hold have more to do with your own personal character than the experiences that may have given rise to them. (reverse)
32. It is a noble thing when someone holds the same beliefs as their parents. (reverse)
33. One should disregard evidence that conflicts with your established beliefs. (reverse)
34. Someone who attacks my beliefs is not insulting me personally.
35. It is important to persevere in your beliefs even when evidence is brought to bear against them. (reverse)
36. Certain beliefs are just too important to abandon no matter how good a case can be made against them. (reverse)
37. Abandoning a previous belief is a sign of strong character
38. Beliefs should always be revised in response to new information or evidence.
39. It makes me happy and proud when someone famous holds the same beliefs that I do. (reverse)

Counterfactual Thinking Subscale

40. My beliefs would not have been very different if I had been raised by a different set of parents.
41. If my environment (family, neighborhood, schools) had been different, I probably would have the same religious views.

APPENDIX D

LOGICAL REASONING ABILITY QUESTIONS

Q1:

An Argument:

The law firm of Sutherlin, Pérez, and Associates is one of the most successful law firms whose primary specialization is in criminal defense cases. In fact, the firm has a better than 90 percent acquittal rate in such cases. Dalton is an attorney whose primary specialization is in divorce cases, so Dalton certainly cannot be a member of Sutherlin, Pérez, and Associates.

The reasoning in the argument is *flawed* because the argument:

a) offers in support of its conclusion pieces of evidence that are mutually contradictory.

b) overlooks the possibility that a person can practice law without being a member of a law firm.

c) concludes that someone is not a member of a group on the grounds that that person does not have a characteristic that the group as a whole has

d) takes a high rate of success among the members of a group to indicate that the successes are evenly spread among the members.

e) states a generalization based on a selection that is not representative of the group about which the generalization is supposed to hold true.

Q2:

Advertisement:

A leading economist has determined that among people who used computers at their place of employment last year, those who also owned portable ("laptop") computers earned 25 percent more on average than those who did not. It is obvious from this that owning a laptop computer led to a higher-paying job.

Which one of the following identifies a *reasoning error* in the advertisement?

a) It attempts to support a sweeping generalization on the basis of information about only a small number of individuals.

b) Its conclusion merely restates a claim made earlier in the argument.

c) It concludes that one thing was caused by another although the evidence given is consistent with the first thing's having caused the second.

d) It offers information as support for a conclusion when that information actually shows that the conclusion is false.

e) It uncritically projects currently existing trends indefinitely into the future.

Q3:

A Decision:

For every 50 dogs that contract a certain disease, one will die from it. A vaccine exists that is virtually 100 percent effective in preventing this disease. Since the risk of death from complications of vaccination is one death per 5,000 vaccinations, it is therefore safer for a dog to receive the vaccine than not to receive it.

Which one of the following would it be ***most helpful*** to know in order to evaluate the argument?

a) The total number of dogs that die each year from all causes taken together.

b) Whether the vaccine is effective against the disease in household pets other than dogs.

c) The number of dogs that die each year from diseases other than the disease in question.

d) The likelihood that a dog will contract another disease such as rabies.

e) The likelihood that an unvaccinated dog will contract the disease in question.

Q4:

Nearly all mail that is correctly addressed arrives at its destination within two business days of being sent. In fact, correctly addressed mail takes longer than this only when it is damaged in transit. Overall, however, most mail arrives three business days or more after being sent.

If the statements above are true, which one of the following must be true?

a) A large proportion of the mail that is correctly addressed is damaged in transit.

b) No incorrectly addressed mail arrives within two business days of being sent.

c) Most mail that arrives within two business days of being sent is correctly addressed.

d) A large proportion of mail is incorrectly addressed.

e) More mail arrives within two business days of being sent than arrives between two and three business days after being sent.

Q5:

Using fossil energy more efficiently is in the interest of the nation and the global environment, but major improvements are unlikely unless proposed government standards are implemented to eliminate products or practices that are among the least efficient in their class.

Objection: Decisions on energy use are best left to the operation of the market.

Which one of the following, if true, most directly undermines the objection above?

a) It would be unrealistic to expect society to make the changes necessary to achieve maximum energy efficiency all at once.

b) There are products, such as automobiles, that consume energy at a sufficient rate that persons who purchase and use them will become conscious of any unusual energy inefficiency in comparison with other products in the same class.

c) Whenever a new mode of generating energy, such as a new fuel, is introduced, a number of support systems, such as a fuel-distribution system, must be created or adapted.

d) When energy prices rise, consumers of energy tend to look for new ways to increase energy efficiency, such as by adding insulation to their houses.

e) Often the purchaser of a product, such as a landlord buying an appliance, chooses on the basis of purchase price because the purchaser is not the person who will pay for energy used by the product.

Q6:

An Argument:

Opponents of allowing triple-trailer trucks to use the national highway system are wrong in claiming that these trucks are more dangerous than other commercial vehicles. In the western part of the country, in areas where triple-trailers are now permitted on some highways, for these vehicles the rate of road accident fatalities per mile of travel is lower than the national rate for other types of commercial vehicles. Clearly, triple trailers are safer than other commercial vehicles.

Which one of the following, if true, most substantially *weakens* the argument?

a) It takes two smaller semitrailers to haul as much weight as a single triple-trailer can.

b) Highways in the sparsely populated West are much less heavily traveled and consequently are far safer than highways in the national system as a whole.

c) Opponents of the triple-trailers also once opposed the shorter twin-trailers, which are now common on the nation's highways.

d) In areas where the triple-trailers are permitted, drivers need a special license to operate them.

e) For triple-trailers the rate of road accident fatalities per mile of travel was higher last year than in the two previous years.