

Individual Differences Influence the Degree of Source Expertise Bias in Syllogistic Reasoning

by

Lisa Boucher  
B.A.H. Psychology, Carleton University, 2011

A thesis submitted to  
the Faculty of Graduate and Postdoctoral Affairs  
in partial fulfillment of the requirements for the degree of

Master of Arts

in

Psychology

Carleton University  
Ottawa, Canada

©2014  
Lisa Boucher

## Abstract

Similar to Copeland, Gunawan, and Bies-Hernandez (2011), the present study investigated the influence of source expertise characteristics on people's susceptibility to bias in syllogistic reasoning. Each syllogistic conclusion was presented by one of two sources: an expert or a non-expert. In the present experiment, the medium of presenting the source information was also altered to include visual images of the sources. Furthermore, individual differences measures were included in order to investigate whether working memory capacity and certain thinking styles (i.e., reflective or actively open-minded thinking) are related to the source expertise bias. In contrast to the source expertise bias that Copeland et al. (2011) had found among their full sample of participants, only a specific sub-sample of the participants in the present experiment exhibited the bias. Namely, people who scored low on the Cognitive Reflection Test (Frederick, 2005) showed the source expertise bias, whereas those with high scores did not.

### Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. Guy Lacroix, for the incredible opportunity and extensive guidance throughout my studies. I would also like to thank my partner, Colin McCulloch, as well as my parents, Steve and Mary, and my brother Jon, for their unremitting and tremendous encouragement. In addition, I must thank my lab mates, in particular Jordan and Georgina, for vastly enriching my academic experience.

## Table of Contents

|   |     |
|---|-----|
| Abstract.....   | ii  |
| Acknowledgements.....                                     | iii |
| List of Tables.....                                       | v   |
| List of Illustrations.....                                | vi  |
| List of Appendices.....                                   | vii |
| Syllogistic Reasoning.....                                | 5   |
| Reasoning Strategies and Mental Models.....               | 8   |
| Belief Bias Effects.....                                  | 11  |
| General Accounts of Persuasion.....                       | 12  |
| Source Credibility Effects.....                           | 15  |
| Relating Syllogistic Reasoning to Source Credibility..... | 20  |
| Individual Differences in Cognition.....                  | 25  |
| Working Memory and Reasoning Ability.....                 | 29  |
| Thinking Dispositions and Reasoning Ability.....          | 34  |
| Experiment.....   | 39  |
| Method.....   | 41  |
| <i>Participants</i> .....                                 | 41  |
| <i>Stimuli and Measures</i> .....                         | 41  |
| <i>Procedure</i> .....                                    | 49  |
| Results.....  | 50  |
| Discussion.....   | 63  |
| References.....   | 80  |

List of Tables

|              |    |
|--------------|----|
| Table 1..... | 42 |
| Table 2..... | 49 |
| Table 3..... | 55 |
| Table 4..... | 60 |

## List of Illustrations

|               |    |
|---------------|----|
| Figure 1..... | 10 |
| Figure 2..... | 28 |
| Figure 3..... | 44 |
| Figure 4..... | 52 |
| Figure 5..... | 52 |
| Figure 6..... | 59 |

List of Appendices

|                 |    |
|-----------------|----|
| Appendix A..... | 92 |
| Appendix B..... | 93 |
| Appendix C..... | 95 |

### Individual Differences Influence the Degree of Source Expertise Bias in Syllogistic Reasoning

Every day, people are presented claims from a variety of sources and they are at times compelled to assess their truthfulness or their validity. Examples abound. For instance, a mother might ask herself if Nutella, a chocolate and hazelnut morning toast spread, should truly be part of a “nutritious” and “healthy” breakfast as used to be claimed by the maker until recently (Thanh Ha, 2012). Parents might worry about giving their children the measles, mumps, and rubella (MMR) vaccine for fear that it will cause autism (for an excellent discussion which debunks this myth, see Goldacre, 2010). Finally, watching late night television, an undergraduate psychology student might seriously doubt an advertiser’s claim that humans only use 10% of their brain “power” (clearly, another debunked myth, see Lilienfeld, Lynn, Ruscio, & Beyerstein, 2010).

Naturally, the extent to which people will exert effort in understanding and assessing the supporting arguments for any given claim will vary depending on personal relevance and importance. Nevertheless, in most cases, people will not be in a position to fully analyze all the information at hand because there is simply not enough time to do so considering all the other tasks that must be accomplished on a daily basis (Simon, 1956). Hence, they will typically rely on cognitive shortcuts known as heuristics, which require little information, are quick to execute, and typically yield reasonably accurate results (Gigerenzer & Brighton, 2009; Kahneman, 2011).

Since the advent of Tversky and Kahneman’s (1974) program of research on cognitive heuristics and biases, psychological research has investigated extensively the tendencies that people exhibit while making judgments in reasoning tasks (Gilovich & Griffin, 2002; Kahneman & Frederick, 2005). Many common errors that people make in situations of uncertainty have been identified. This is because heuristics are often employed when people’s understanding

regarding the relevance of information is lacking. That is, if people are familiar with a topic of information, they are likely to base decisions on what they know about the topic. In contrast, if people are not familiar with a topic, it is likely they will rely on other features of the information that they do understand (Bromme, Kienhues, & Porsch, 2011; Weisberg, Keil, Goodstein, Rawson, & Gray, 2008). Such reliance can help or hinder their decision-making. Moreover, certain individual differences in human cognition are also related to how likely it is that a given person will rely on heuristic-based strategies. Cognitive differences among people influence how well they are able to engage in effortful analysis if necessary, or how likely they are to disengage from such critical processing and instead rely on less effortful, “rule of thumb” strategies (Stanovich & West, 1997).

Formal reasoning tasks are particularly useful for studying people's use of heuristics. They require people to focus solely on the logical structure of the task and ignore contextual information in order to reason accurately. People encounter this type of logical reasoning in different real-life circumstances including academic instruction in abstract logic and mathematics, standardized examinations (e.g., SAT or GRE), and of course laboratory experiments in introductory psychology courses at universities (Khemlani & Johnson-Laird, 2012). One type of formal reasoning task that has received considerable attention in cognitive psychology is syllogistic reasoning.

Syllogistic reasoning problems consist of two premises and a conclusion. For any syllogism to be considered valid, the conclusion *must* follow logically from the premises. Syllogisms have been very useful to assess the conditions that encourage people to use certain strategies for reasoning, as well as the conditions that result in people making particular errors. Because people find formal reasoning difficult, they often use heuristic-based strategies. For

example, the “atmosphere” of syllogistic premises has been shown to affect people’s reasoning (Woodworth & Sells, 1935). In general, people are more likely to think a conclusion is valid if the quantifier has the same atmosphere (affirmative, negative, universal, or particular) as one or both of the quantifiers in the premises. Another classic heuristic associated with syllogistic reasoning is known as the *belief-bias effect* (Evans, Barston, & Pollard, 1983). It occurs when participants fail to correctly determine the validity of a conclusion using logic because doing so contradicts their world knowledge (e.g., concluding that “smoking is good for the health”).

In many real-life reasoning situations, one aspect of information that people often depend on is their knowledge about the credibility of the source providing the information (Bromme et al., 2011). For example, if a patient receives the diagnosis that she has a rare disease which she had not previously heard of, it is more likely that she will depend on advice from a medical doctor than it is that she will depend on advice from anyone else. Still, credibility information about a source is not always going to be a reliable indicator of whether a person should believe a piece of information. Using source credibility information to make decisions might be a good strategy if very little is known about the topic or if the proper decision-making process required to execute a given task is unavailable. Yet, with access to other informative indicators (e.g., a topic-related inconsistency or a requirement to ignore source credibility), people might still be able to determine that a usually reliable source is not correct in a particular situation.

The present thesis investigated the influence of source credibility on syllogistic reasoning. The expertise of a source of information, in particular, is one of the most well-studied source characteristics in social psychology paradigms. Copeland, Gunawan, and Bies-Hernandez (2011) recently combined these two primarily distinct research areas. These researchers determined that source credibility effects were quite similar to belief-bias effects because both

involved the influence of people's prior beliefs. The finding that source credibility effects occur in a syllogistic reasoning paradigm is important because source credibility has often had only a peripheral influence on people's decision-making, secondary to argument quality, in the attitude change literature (Petty & Wegener, 1998; Sparks & Rapp, 2011). Moreover, people should ignore source credibility information in syllogistic reasoning tasks because it is irrelevant to solving syllogisms accurately. In contrast, the relevancy of source information is likely to be less clear in attitude change tasks because the questions often directly ask for participants' opinions. Hence, it is sensible to expect that source credibility should have a smaller effect on people's syllogistic reasoning ability in comparison to its effects in attitude change tasks. However, Copeland et al.'s findings contradicted such an intuition. It is clear that more study of source credibility biases in syllogistic reasoning will provide a clearer estimate of the strength of these biases. Furthermore, an investigation into how formal reasoning ability is influenced by differences in people's cognitive processes is warranted in order to help explain why the biases occur.

Accordingly, my thesis had several goals. First, I attempted to replicate Copeland et al.'s (2011) finding that source credibility (and more specifically, source expertise) influences people's formal reasoning abilities. Second, I extended Copeland et al.'s methodology to examine the magnitude of source expertise on reasoning when source information was visually depicted (i.e., an image of the expert is presented) rather than simply written. Third, my thesis used both accuracy and response time measurements to investigate the relationships among individual differences in people's cognition and source credibility effects.

To reach these goals, I proceeded as follows. First, I present a detailed discussion of formal syllogistic reasoning, describe research on the use of strategies in formal reasoning

paradigms, and on the influence of prior beliefs on reasoning accuracy. Subsequently, I discuss persuasion and source credibility research that pertains to expertise. Third, I thoroughly describe Copeland et al.'s (2011) experiment. Fourth, I outline the individual differences literature relating working memory, thinking dispositions and reasoning ability. Finally, the goals and hypotheses, methodology, results, and discussion for my thesis experiment are presented.

### **Syllogistic Reasoning**

Syllogisms are deductive statements comprising three parts: two premises and a conclusion. Each premise relates a single “middle term” with one of two “end terms” while the conclusion relates the two end terms (e.g., All A are B. No B are C. No A are C.). Four “figures” are constituted through the possible presentation orders of middle and end terms in the two premises (1 = A-B, B-C; 2 = B-A, C-B; 3 = A-B, C-B; and 4 = B-A, B-C). The determiners *all*, *some*, *no*, and *some not* constitute the four syllogistic “forms” (or moods). These moods range from affirmative (all, some) to negative (no, some not), and universal (all, no) to existential (some, some not).

The standard set of syllogisms that researchers typically use comprises 64 different logical structures. This set represents the complete variation of mood in each of the first two premises as well as the figure of each pair of premises (i.e., 4 moods in premise one X 4 moods in premise two X 4 figures; Khemlani & Johnson-Laird, 2012). Hence, these 64 syllogisms include all combinations of figure and form in a syllogism, with the exception of the form and order of the terms in the conclusive statement. Note that although some researchers do alter the order of the terms (between A-C and C-A) in syllogistic conclusions, they do not typically include each of the 64 pairs of premises more than once in their studies. That is, one half of the 64 syllogisms might be given A-C conclusions, while the other half would be given C-A

conclusions. Furthermore, it should be noted that technically it is possible to consider syllogistic argumentation according to a total of 256 or 512 distinct syllogisms. That is, if all possible combinations of conclusive statements were also included (64 X 2 conclusion term orders X 4 conclusion moods), then each pair of premises would be considered eight times (once with each of eight different potential conclusions). However, these larger syllogistic sets mainly introduce redundancy because only the two premises in a syllogism are needed to establish whether a logically valid conclusion is possible or not (Khemlani & Johnson-Laird, 2012).

Syllogisms have often been used to frame tasks in studies containing logical arguments with determiners. Researchers in these studies often give participants explicit instructions to ignore their prior knowledge and instead to focus on the logical structure of the syllogisms. These tasks typically require participants to make a decision response of either logical validity (representing necessity) or logical invalidity (the presence of non-necessity). Hence, “valid” syllogisms have at least one conclusion that establishes a definite relationship between the end terms in consideration of all possible interpretations of the premises (i.e., they have necessity; Khemlani & Johnson-Laird, 2012). Considering the 64 syllogism set, there are 27 valid syllogisms. The remaining 37 syllogisms are considered “invalid” due to having no conclusions that hold in all possible relations among the premises. Furthermore, most experiments including syllogistic reasoning tasks employ one of two general procedure types: evaluation of given conclusions or production of a conclusion. Both tasks yield similar results (Johnson-Laird & Bara, 1984).

Psychological research has determined that people exhibit several consistent, robust patterns of results in syllogistic reasoning experiments. On the whole, there is a strong tendency for people to accept syllogisms as valid, making accuracy higher for valid conclusions than for

invalid conclusions. Moreover, the figure of the syllogism has an impact on the difficulty level of reasoning with particular syllogisms. For instance, some figures lead participants to prefer A-C order conclusions whereas others lead them to prefer the C-A order. When these preferences conflict with the logical validity of the syllogisms, participants' performance declines (Khemlani & Johnson-Laird, 2012). Mood effects also influence the difficulty level of syllogisms. People reason more accurately on syllogisms containing positive and universal moods (Khemlani & Johnson-Laird, 2012; Woodworth & Sells, 1935). Specifically, they are more accurate on syllogisms that contain the mood *all* rather than syllogisms that contain *some* or *no*. They have the most difficulty reasoning with syllogisms containing *some not*. On the whole, the level of difficulty of syllogisms varies immensely. Whereas most people are able to produce accurate conclusions to certain syllogisms (e.g., All A are B. All B are C. All A are C.), very few people are able to produce accurate conclusions for others (e.g., No A are B. All B are C. Some C are not A.).

In most studies, participants are instructed to indicate that a syllogism is valid based on a conclusion that holds in all possible circumstances relating the premises. People are also able to distinguish between possible and impossible syllogisms. A possible syllogism has a conclusion that establishes a relation among the premises in at least one interpretation. In contrast, an impossible syllogism has no conclusions that establish a relation among the premises in any possible interpretations. Evans, Handley, Harper, and Johnson-Laird (1999) conducted research to better understand people's reasoning tendencies if given instructions to categorize syllogisms based on possibility rather than necessity. Using this method, they found that more syllogisms were accepted under possibility instructions than under necessity instructions. Moreover, they found that people consistently tended to classify certain possible syllogisms as valid while

classifying other possible syllogisms as invalid. Hence, such syllogisms might easily be divided into “possible strong” and “possible weak”: two categories generated according to whether the majority of people consider the syllogisms valid versus invalid, respectively. The pattern of findings also showed that when comparing possibility instructions to necessity instructions, the difference in the acceptance of syllogisms was greater for possible syllogisms than it was for necessary or impossible syllogisms. Furthermore, evidence that possible strong syllogisms are very likely to be incorrectly classified as valid, if given necessity instructions, agrees with the fact that greater reasoning biases are typically evident for invalid syllogisms rather than valid syllogisms. Evans et al.’s (1999) classification of syllogism types was integrated into my thesis’s methodology.

### **Reasoning Strategies and Mental Models**

There have been many theories and models of the formal reasoning process over the course of its psychological study. These theories have attempted to explain the strategies, heuristics, and tendencies that people exhibit during syllogistic reasoning tasks. In an extensive meta-analysis of the present state of explanation in the syllogistic reasoning literature, Khemlani and Johnson-Laird (2012) critically outlined twelve different models and put several datasets to a stringent test. They divided these theories into three broad categories in accordance with the types of strategies that are outlined in the theories. Firstly, there are formal rule-based strategies, which explain people’s reasoning through aspects of basic logical rules, such as the processes of deduction, verbal substitution procedures, or the properties of quantifiers. These include PSYCOP (Braine & Rumin, 1983; Rips, 1994), verbal substitutions (Ford, 1995), and monotonicity (Geurts, 2003). Secondly, there are heuristic-based strategies, including atmosphere (Begg & Denny, 1969; Woodworth & Sells, 1935), matching (Wetherick &

Gilhooley, 1995), illicit conversion (Chapman & Chapman, 1959; Revlis, 1975), and the probability heuristics model (Chater & Oaksford, 1999). As noted previously, heuristics are accessible, “rule of thumb” type strategies which typically explain people’s reasoning errors as departures from deductive reasoning. Thirdly, there are strategies that make use of diagrams or sets, such as mental models (Johnson-Laird, 1975; 1983), Euler circles (Erickson, 1974), Venn diagrams (Newell, 1981), verbal models (Polk & Newell, 1995), and source-founding (Stenning & Yule, 1997). Two of these theories propose that people create visual representations of a specified type of diagram (i.e., Euler circles or Venn diagrams). Similarly, the ideas underlying the theories of mental models and verbal models both focus on the manipulation of mental representations. Verbal models, however, focus more on the processes of encoding and re-encoding, whereas mental models focus more on the search for alternative models (and the use of such counterexamples for falsification). Finally, the source-founding theory does not propose that people’s mental representations follow a particular modality. Rather, it focuses on the idea that people tend to identify a “source premise” on which to found their conclusions.

The mental models account is prominent in the literature (Johnson-Laird, 1983; Johnson-Laird & Bara, 1984). It suggests that participants use a strategy of forming one or more mental representations of a syllogism that demonstrate the conclusion as either valid or invalid. The model predicts that syllogisms requiring a larger number of mental representations of the premises will cause more difficulty during reasoning than syllogisms requiring a single or smaller number of mental representations. For instance, consider the following syllogism: All A are B. All B are C. All A are C. The relations between the premises require people to consider only a single mental model which shows the conclusion as valid (see Figure 1). Yet consider another syllogism: Some A are B. Some B are C. Some A are C. The relations between these

premises elicit two possible interpretations that people must consider in order to reason accurately. That is, in one mental model, the parts of A and C that are also B are overlapping whereas in a second mental model those parts are not overlapping. Hence, depending on the mental model that is generated, people may rightly conclude that the syllogism is invalid or wrongly conclude that it is valid. Note that each syllogistic structure has a determined number of possible mental representations.

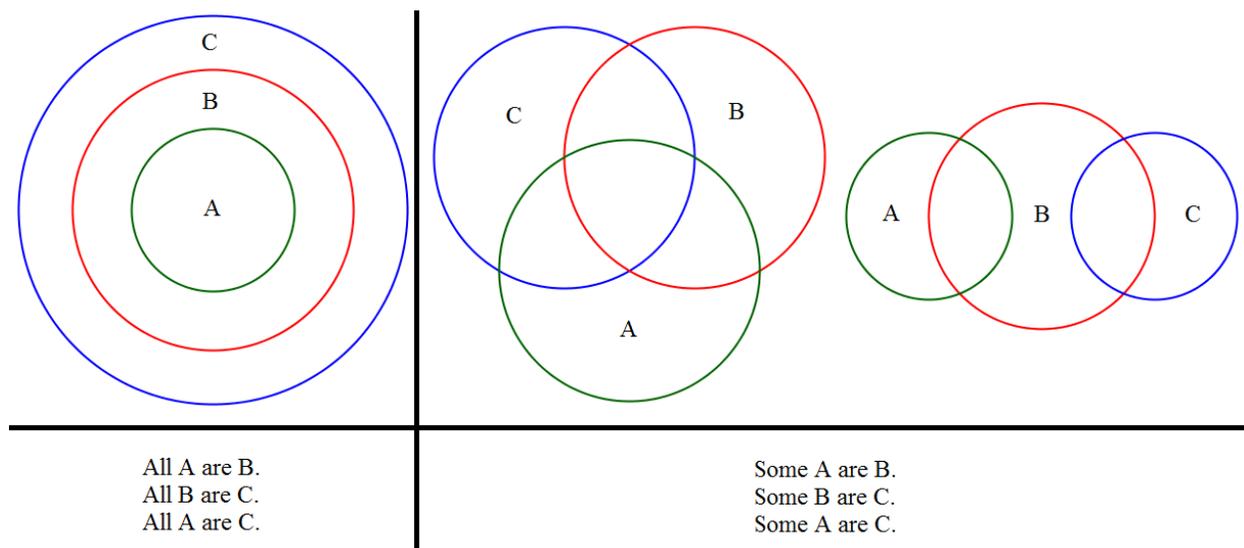


Figure 1. Visual depiction of the text example (using Euler circles) which compares a syllogism that elicits one mental model (left) to a syllogism requiring two mental models (right).

While no single theory can successfully predict all the extant syllogistic reasoning data, it appears that the mental models approach is the one that has received the most empirical support (Khemlani & Johnson-Laird, 2012). Khemlani and Johnson-Laird's (2012) meta-analysis explored only 7 of the 12 theories because the others were incomplete theories (e.g., they did not include specific predictions). They found that the mental model account could explain syllogistic reasoning accuracy as well or better than the strongest competing theories (i.e., verbal models and conversion). Notwithstanding these theoretical considerations, mental models can serve as a metric for a syllogism's level of difficulty; as the number of mental models that a syllogism

yields increases, so does its relative difficulty. The relationship between mental models and people's individual differences in regard to working memory is discussed later.

### **Belief-Bias Effects**

People consistently show certain biases when they evaluate the validity of syllogisms. Arguably, the most extensively studied one is the belief-bias effect. It describes an individual's tendency to accept believable conclusions and reject non-believable conclusions regardless of the validity of the conclusions (Ball, Phillips, Wade, & Quayle, 2006; Evans et al., 1983; Markovits & Nantel, 1989; Newstead, Pollard, Evans, & Allen, 1992). In studies of people's belief-based biases, syllogistic content that varies in "believability" is typically used in place of the abstract arguments and conclusions (i.e., A, B, and C) that are used by logicians. The inclusion of such content that differs according to how well it conforms to people's prior knowledge (i.e., believability) has been shown to elicit belief-bias effects (Evans et al., 1983). Even though content is irrelevant to a syllogism's logical validity, the interference from people's prior knowledge regarding the content causes reasoning errors. For instance, an invalid syllogism with a believable conclusion (e.g., Some vehicles are aircraft. Some aircraft are helicopters. Therefore, some vehicles are helicopters.) is more likely to be accepted than an invalid syllogism with an unbelievable conclusion (e.g., Some cats are predators. Some predators are reptiles. Therefore, some cats are reptiles.). Moreover, logic and belief interact so that these biases typically influence responses to invalid syllogisms to a greater degree than responses to valid syllogisms (Ball et al., 2006; Evans et al., 1983).

Despite the existence of belief-bias effects, any type of meaningful content is generally known to increase people's ability to reason accurately because it is more likely to be easily encoded (Khemlani & Johnson-Laird, 2012). Hence, because many people have difficulty

reasoning with syllogisms overall, the use of neutral content instead of letters or meaningless content is likely to make a syllogistic reasoning task easier for people to complete accurately. Neutral content consists of arguments and conclusions that are neither intrinsically believable nor unbelievable (i.e., it is equally plausible that they will be considered either believable or unbelievable). Therefore, if investigating how syllogistic reasoning ability relates to factors that are extraneous to the syllogistic content, the use of neutral content allows researchers to control for potential belief biases that distort accuracy performance. This is the type of syllogistic reasoning task design that was used in my thesis following the method of Copeland et al. (2011). The information within the syllogisms was neutral in order to measure any biases resulting solely from information that was provided extraneous to the syllogisms (i.e., the information pertaining to source credibility).

### **General Accounts of Persuasion**

Before turning to research that has focused specifically on the relation between syllogistic reasoning and source credibility, I first outline several prominent models in the attitude change and persuasion literature that are crucial to consider with regard to the study of source credibility. Subsequently, I discuss a variety of important effects in the source credibility literature that are relevant to the present thesis.

Early research on persuasion provided the field with an array of complex findings (Petty & Briñol, 2012; Chaiken & Ledgerwood, 2012), which made generalization difficult and left the topic area with a lack of overarching theory. Nevertheless, two similar theories eventually emerged. Both postulated that people use two different kinds of cognitive processes when they consider information in situations of persuasion. These theories are known as the Elaboration

Likelihood Model (ELM; Petty, 1977; Petty & Cacioppo, 1981) and the Heuristic Systematic Model (HSM; Chaiken, 1978; Chaiken, Liberman, & Eagly, 1989).

According to the ELM, there are two prominent routes leading to the occurrence of persuasion: the *central route* involving a high degree of thought and the *peripheral route* involving a low degree of thought (Petty & Briñol, 2012). Processes of attitude change may operate anywhere along this continuum from low to high degrees of thought. Judgements or attitudes that are based on processes of high thought have been shown to persist longer over time and to have more consequences for behaviour (Petty, Haugtvedt, & Smith, 1995). Situational and individual difference variables influence the route of persuasion by moving people back and forth on the thinking continuum. Although the influence of a given variable is relatively predictable if people use the peripheral route, it is less predictable if people use the central route because the variable may or may not be deemed relevant. Moreover, any variable that alters the amount of thinking an individual does is likely to change the nature of their information processing. For example, Petty, Cacioppo, and Goldman (1981) manipulated the degree of personal relevance that a situation had for individuals (i.e., by either telling them that a policy outlined in an argument would or would not be implemented at their university). Then, they measured the influence of argument quality (weak or strong) and source expertise (expert or non-expert) on the amount of persuasion people experienced in favour of the argument. Their results indicated that people relied on different types of information depending on how personally relevant the situation was for them. When the relevance conditions were high, people used the quality information to judge the argument. That is, the argument was judged as better if it had strong quality than if it had weak quality. In contrast, under low relevance conditions, people

used the source information to judge the argument. Hence, the argument was judged as better if it was provided by the expert than if it was provided by the non-expert.

The HSM is also a dual-system approach that presents an explanation for attitude change. It postulates two types of processing – *systematic* and *heuristic* – which necessitate high and low degrees of mental effort, respectively (Chaiken & Ledgerwood, 2012). The HSM hypothesizes that people must be able *and* motivated to engage in systematic processing. In contrast, heuristic processing relies on readily available and easily understood cues that may be conscious or unconscious. The main determinant of how much effort people exert in a situation of persuasion is the size of the gap between their actual confidence and their desired confidence. Thus, people initially might rely on heuristic cues if the gap is small, but if the gap is larger, they will be more likely to exert additional mental effort in order to increase their confidence. The HSM also takes into account the fact that people have multiple motives during decision-making (i.e., accuracy motivation as an objective concern, and defense and impression motivations as biased concerns). Finally, the model indicates that systematic and heuristic processing can both co-occur and interact.

Many findings in support of both the ELM and the HSM indicate that the motivation to accurately assess information in situations of persuasion depends on the degree of personal relevance, consequences, or involvement in a given situation (Chaiken, 1980; Petty & Cacioppo, 1979; Petty et al., 1981). Most of these studies, however, have been conducted from a social psychological perspective and thus have used situational manipulations to adjust people's level of motivation. Nonetheless, similar findings would be expected in studies that measure cognitive individual differences which might be relevant to people's dispositional level of motivation. For instance, people who score low on a measure of their tendency to think carefully about

information might also have lower motivation because they believe that many tasks do not have personal relevance to them. Accordingly, people who score highly on this type of thinking measure might also have higher motivation because they consider more tasks to be personally relevant to their lives. Hence, in regard to the influence of source credibility cues versus the quality of an argument, it is expected that differences in people's dispositional traits also influence their degree of persuasion. Cacioppo, Petty, and Morris (1983) have found that individual differences in Need for Cognition (i.e., the tendency for people to engage in thinking) influence people's judgements. Thus, if all people experience the same given situation and they still show differences in their judgements, the resulting explanation should consider dispositional causes.

### **Source Credibility Effects**

Source credibility is the degree to which a provider of information (i.e., the source) can be relied upon to be accurate (i.e., is credible). A majority of the research on source credibility has come from social psychology and it has made use of persuasion, attitude change, and consumer choice paradigms. On the whole, characteristics of trustworthiness and expertise tend to have the greatest influence on people's acceptance of information (Pornpitakpan, 2004). Hence, more credible sources are generally considered to be more trustworthy and to have more expertise. Still, a source's likeableness, attractiveness, personality, and power have also been found to influence people's decision making (Demeure, Bonnefon, & Raufaste, 2009; Kilpatrick, Manktelow, & Over, 2007; Maddux & Rogers, 1980; Raven & French, 1958; Reinhard, Messner, & Sporer, 2006). Moreover, the influence of a source's expertise on people's decision-making is relative to a person's expectations about how their own level of expertise compares to the source's level of expertise in regard to the topic at hand (French, Garry, & Mori, 2011). That

is, if expertise levels are similar then the source is not likely to influence a person's decisions, whereas if the source has greater expertise than the person then it is more likely to influence decisions. In addition, as mentioned previously, people are more likely to be influenced by source credibility during tasks of low importance compared to tasks of high importance (Chaiken & Maheswaran, 1994; Petty & Cacioppo, 1986). Hence, the degree of motivation that people have to perform accurately in a task is expected to have a critical impact on their susceptibility to source credibility effects. Studies have found source credibility effects in a variety of tasks, including text comprehension (Sparks & Rapp, 2011), attitude judgement (Chaiken & Maheswaran, 1994), behavioural intent judgement (Homer & Kahle, 1990), and even information recognition in circumstances involving misinformation (French et al., 2011). In addition, several reasoning tasks involving conditional deductions (Stevenson & Over, 2001; Demeure et al., 2009) and deontic inferences (Kilpatrick et al., 2007) have revealed source credibility effects.

Effects stemming from a source's expertise, in particular, have been the focus of extensive research. It is likely that expertise is the most dominant dimension of a source. Yet, expertise has numerous potential influences that depend on how it interacts with other source characteristics and contextual factors (Homer & Kahle, 1990; Wilson & Sherrell, 1993). Recall that people are more likely to defer to expertise if the topic under scrutiny is of low personal relevance to them (Hovland & Weiss, 1951; Petty, Cacioppo, & Goldman, 1981). Thus, source expertise effects should occur more often if people are evaluating neutral information rather than information that they have strong beliefs about. Furthermore, expertise has a greater effect when people are distracted (Kiesler & Mathog, 1968) and when information is paced externally rather than self-paced (Andreoli & Worchel, 1978). Additionally, expertise might even influence

reasoning in contexts in which no source information is explicitly provided. For instance, Weisberg et al. (2008) found that people accept both good and bad explanations of psychological phenomena more often if they include neuroscience information. The statements that they used did not explicitly include any description about the source of the information, but people might have assumed that the information was provided by neuroscientists (Schoenherr, Thomson, & Davies, 2011). Hence, because participants had limited experience analyzing neuroscientific knowledge, an expert in that field would be considered an authoritative source. Thus, the assumption would have increased participants' acceptance of neuroscience arguments. My contention here is simply that source effects regarding expertise might be more prevalent than is typically recognized.

Nonetheless, source credibility has often been found to have a limited impact on people's judgement of statements (Petty & Wegener, 1998; Sparks & Rapp, 2011). Because the quality of an argument is typically a more central factor regarding how people judge information, credibility characteristics of the source of an argument do not always emerge as an important factor. An example study will help illustrate the methodological challenges associated with showing the role of source credibility on people's evaluation of information. Sparks and Rapp (2011) conducted four experiments investigating the effects of source credibility on participants' expectations of the outcomes for characters in texts they read. Each text involved a meeting between a reporter and one of two informants (who was presented as either a reliable or unreliable source). The texts were written in a conversational transcript style, in which stories about various characters' actions were presented. During the stories, these characters were described as having specific personality traits, such as "cheater" or "generous". Moreover, a character's actions at the outcome of the story were designed to be either consistent or

inconsistent with the character's traits. Throughout the text reading process, participants' reading latencies were measured for each sentence. This method was used to determine if source credibility had an effect on people's expectations for the character's outcome behaviour. That is, if participants took longer to read trait-inconsistent outcomes from the credible informant or trait-consistent outcomes from the non-credible informant, then it could be concluded that source credibility was influencing participants' ability to process the information. In their four experiments, Sparks and Rapp progressively increased the amount of source credibility information that they gave their participants. The latter experiments also encouraged the participants explicitly to pay attention to the source credibility information. Despite these manipulations, no significant effects of credibility on text evaluations were found until the last experiment, in which participants were explicitly instructed to consider the source credibility information and then to make likelihood judgements regarding the text outcomes. Overall, the coherence of the text information (i.e., trait consistency of the characters) was the chief influence on people's evaluations. Hence, it is clear that people pay little attention to source credibility in some situations.

Regardless of such weak influences resulting from source credibility in certain paradigms, others have been more convincing. For instance, Stevenson and Over (2001) conducted several experiments to investigate the effect of source expertise on people's level of uncertainty regarding outcomes. They presented an initial rule statement which was followed by two assertions: one speaker provided information pertaining to the rule and then a subsequent speaker denied it (in order to portray a conversational context). Overall, Stevenson and Over determined that people were more certain about conclusions if an expert corrected a novice's statement rather than the opposite. It is not surprising that they found such source expertise

effects because their experimental instructions explicitly encouraged participants to “consider carefully” the sources of the statements. Furthermore, their conditional reasoning design included contradictory statements given in accordance with a rule. Hence, this procedure also gave implicit encouragement for participants to take the expertise characteristics into account while choosing among alternative outcomes. Such implicit encouragement results from the fact that the “correct” responses in this type of experimental design are often ambiguous (Chaiken & Maheswaran, 1994). That is, although certain answers might be considered the *most* correct, the task requirements allow for flexibility in interpreting which of the various responses are more or less correct. Thus, such an informal reasoning task likely leads to the impression that it is acceptable, and even sensible, for participants to use source information to aid them in reasoning. Furthermore, the extent to which additional factors might influence people’s likelihood to use source information is difficult to control through these experiments.

Although many studies that find source credibility effects do not explicitly require participants to consider the source information during decision-making, there is generally some degree of implicit encouragement. Furthermore, participants might consider many tasks in reasoning studies to be of low importance or relevance in comparison to attitude change studies which sometimes involve explicit belief-persuasion attempts. Such explicit persuasion often results in participants devoting additional consideration to decisions, which then leads them to reason more carefully and to be less influenced by source credibility (Reinhard et al., 2006). Hence, because Stevenson and Over’s (2001) study employed relatively ambiguous correct responses, without involving precise attitude change attempts, the design created an ideal context for people to utilize source credibility information. This use is certainly not advantageous in *all* real-life circumstances, in which the relevance of source credibility varies immensely and is

often unclear (Evans, 2002). Thus, because people's responses in studies involving informal reasoning, such as Stevenson and Over's, are more likely to be affected by contextual factors, these findings are largely unable to account for expertise biases that people might exhibit when source information is actually irrelevant (i.e., in any circumstance in which people must ignore contextual information in order to reason accurately). In contrast, similar experiments using formal reasoning – and more specifically syllogistic reasoning tasks – should provide a clearer picture concerning the strength of source credibility effects. Firstly, it is possible to design a syllogistic reasoning task that will minimize potential belief-bias effects, thus reducing influences from such extraneous factors. Secondly, it is possible to avoid including explicit instructions for participants to consider source credibility information in regard to their reasoning. Furthermore, due to syllogisms having correct answers that are not ambiguous (although see Chater & Oaksford, 1999), implicit encouragement to consider source information is also likely to be reduced. Hence, using a syllogistic reasoning task in my thesis allowed for a controlled investigation into the strength of people's biases regarding source credibility in a context that did not explicitly encourage them to consider the source information.

### **Relating Syllogistic Reasoning to Source Credibility**

Although source credibility effects were displayed in the previously discussed reasoning paradigms, similar findings have not been thoroughly investigated in syllogistic reasoning paradigms (Chaiken & Maheswaran, 1994; Pornpitakpan, 2004). Considering the robustness of certain source biases, such as trustworthiness and expertise, it is not surprising that these characteristics would be applied first within syllogistic reasoning tasks. Yet, to the best of my knowledge, only Copeland et al. (2011) have focally investigated source credibility effects in

syllogistic reasoning. Because my thesis research replicated and extended Copeland et al.'s experimental design, I now describe their method in detail.

Copeland et al. (2011) compared the trustworthiness and expertise of sources in two separate experiments using a standard syllogistic reasoning paradigm. In order to make source credibility the central focus of the study, a cover story was developed and the syllogisms contained neutral content to control for belief-bias effects that might conflict with the expected source effects. The fictional story involved “facts” (representing the first two syllogistic premises) that were supposedly obtained through survey research of a small town’s residents. These facts concerned hobbies and status terms. Participants were told that their task was to evaluate the conclusion that followed each pair of facts. Furthermore, they were instructed that this conclusion was an interpretation provided by one of two different town residents (whose identity was to be presented simultaneously with each syllogism and in a random order). Prior to the reasoning task, these sources of the information were outlined with character sketches, which were adapted from the descriptions of the two informants used in Sparks and Rapp (2011). Each source represented a type of individual: a helpful fireman as *honest* versus a sneaky treasurer as *dishonest* (in Experiment 1); and a distinguished professor as *expert* versus a handy mechanic as *non-expert* (in Experiment 2). Furthermore, Copeland et al. asked participants three questions about the sources immediately after reading each description. These questions involved only basic aspects of the descriptions and it encouraged participants to pay close attention to the source information.

Copeland et al. (2011) employed “necessary” syllogisms to be considered correct and “possible strong” syllogisms to be considered incorrect (see Evans et al., 1999). Recall that possible strong syllogisms are similar to possible weak syllogisms because they have at least one

representation indicating that the syllogism holds true and an alternative representation indicating that it does not hold true. However, people typically conclude that possible strong syllogisms are valid, whereas they typically conclude that possible weak syllogisms are invalid. Impossible syllogisms are those that have no representations in which the conclusion holds true. Possible strong syllogisms were chosen because a pilot study had shown that no source credibility effects were present when impossible syllogisms were used. This result is likely explained by previous findings indicating that participants' experience with a given experimental task (i.e., in which for instance they realize that both sources are equally correct or incorrect) might supplant the influence of any credibility impressions they have of the sources (Gordon & Bryant, 2010). Possible weak syllogisms were not used either in order to ensure that all the conclusions seemed reasonable. Thus, the possible strong syllogism type was used to maximize the likelihood of finding the potential source credibility biases.

Copeland et al.'s (2011) results indicated that source credibility characteristics interacted with logical type of the syllogisms. People were significantly more accurate on the possible strong syllogisms for the dishonest source compared to the honest source. Similarly, people were significantly more accurate on the possible strong syllogisms for the non-expert source compared to the expert source. Hence, participants performed *worse* when they had to evaluate an incorrect syllogism delivered by the honest or expert source rather than the dishonest or non-expert source. A common explanation for this type of phenomenon, in accordance with the critical thinking literature (Sá, Stanovich, & West, 1999), is that people tend to disengage from further thinking if information seems credible initially. The pattern of results is clearly consistent with typical belief-bias effects. Whether the information that elicits people's prior beliefs to influence their reasoning is either extraneous to or contained within syllogisms, the biasing effects tend to

operate in a similar manner. Moreover, source biases are present even in a formal reasoning task in which participants are expected to ignore their prior knowledge and use only the logical structure to make accurate judgements.

Copeland et al. (2011) did not provide explicit instructions to participants to focus on necessity in their evaluations of the syllogisms. Because of the cover story that these authors included, it would have been confusing for participants to be told to ignore the source information and focus only on logical validity (or even to teach them how to do so through a training session). Khemlani and Johnson-Laird (2012) note that even without explicit instructions, participants commonly understand that judgements should be based on the logical structure of the premises and they attempt to respond accordingly. Hence it was not necessary to include such instruction in my thesis, especially because one of the primary goals was to replicate Copeland et al.

Copeland et al.'s (2011) study provides the only focal investigation of source credibility biases in syllogistic reasoning to date. Their experiment, however, only elicited a small effect size for the source credibility by syllogism type interaction. When the partial eta squares that they reported (and that can be susceptible to problems of interpretation, see Levine & Hullett, 2002) are translated into Cohen's  $d$ , they yield a value of approximately .25 (i.e., a small effect). This effect size is not surprising considering Sparks and Rapp's (2011) findings indicating that source credibility effects tend to be difficult to elicit. However, this raises a question regarding the nature of source credibility effects. It may be the case that they are uniformly small; that is, they play a secondary role in participants' evaluations of information and that other factors (e.g., argument quality, character consistency, etc.) are more important. It is also possible, however, that participants might fail to consider source expertise on given trials for different reasons (e.g.,

the retroactive interference created by the consideration of the problem, forgetting the task instructions, etc.). The difficulties that Sparks and Rapp experienced trying to generate a source expertise effect with written materials make both these possibilities plausible. If this hypothesis is correct, one might then conclude that source expertise effects are actually larger than the literature suggests and that previous experimental designs have led to their underestimation.

In an extension of this research topic, my thesis experiment included an additional, more salient method of source presentation in an attempt to increase the effect size of source expertise effects on syllogistic reasoning. That is, the expertise information was presented using a visual medium (i.e., pictures of the experts) as well as the written medium used in Copeland et al. (2011). Although the majority of similar studies have been conducted using written information, the medium of source presentation has been shown to influence the effect size of source credibility's influence on judgements (Wilson & Sherrell, 1993). Moreover, one study had participants rate "how involved they were with the communication" and found that television was considered significantly more involving than audio, which was in turn significantly more involving than the written medium (Andreoli & Worchel, 1978). On the other hand, results from studies comparing three mediums of presenting source information have found that different sources are more persuasive in particular mediums. Namely, trustworthy or likeable sources were more persuasive through video or audio mediums, whereas untrustworthy or unlikeable sources were more persuasive through a written medium (Andreoli & Worchel, 1978; Chaiken and Eagly, 1983). Chaiken and Eagly (1983) referred to a "vividness" phenomenon in regard to the increased influence on persuasion that source information has if it is presented through video and audio mediums rather than written.

Nevertheless, investigation into visual media still seems to be lacking in this research area. Many studies have employed source photos to determine how various specific traits (i.e., trustworthiness, attractiveness, etc.) influence participants' decision-making (Nguyen & Masthoff, 2007; Porter, Brinke, & Gustaw, 2010). Few, however, have compared the influence of providing such photos to the influence of other mediums of presenting source information. Overall, it is likely that photos are considered less salient than videos but more salient than written information. Hence, my thesis employed photos to compare their effect to the influence of the written medium used in Copeland et al. (2011).

### **Individual Differences in Cognition**

People's susceptibility to belief biases and other errors in reasoning have been connected to their individual differences in cognitive processing (Kahneman, 2011; Quayle & Ball, 2000; Stanovich & West, 1997). Accordingly, it is likely that people's individual differences in cognitive processing also relate to the degree to which they are biased by source credibility. Hence, my thesis included several individual difference measurements in order to investigate which cognitive factors lead people to consider or ignore source credibility information while reasoning. In addition, the role of people's individual differences in regard to source credibility's influence on reasoning has important implications for how the effect size of such results should be considered. That is, for example, people with lower levels of cognitive abilities or rigid thinking dispositions might be influenced by source credibility to a large degree (exhibiting a large effect), whereas people with higher levels of cognitive abilities or flexible thinking dispositions might not be influenced at all (exhibiting no effect). Before proceeding to a discussion of specific cognitive processes that are related to reasoning ability, I outline some

major accounts of human cognitive processing which provide explanation for the individual differences that were included in my research.

During the last few decades, many scholars of reasoning have considered dual-process accounts of the mind an effective way to conceptualize how human cognition operates (Evans, 2003; Kahneman, 2011; Norman & Shallice, 1980; Stanovich, 2004; 2009). These dual cognitive operations involve heuristic (“System 1”) and analytic (“System 2”) processes. The processes of System 1 provide mainly automatic responses that are relatively quick. The major domains of System 1 include both processes of intuitive and procedural cognition. In contrast, System 2 processes provide mainly controlled responses, which are more lengthy and effortful compared to System 1 processes. This latter system has generally included both algorithmic and rational levels of analysis. That is, in dual-process models, processes of systematic thinking (algorithmic) and processes of reflective thinking (rational) have been considered the major cognitive functions of System 2. An example will illustrate a typical contrast between operations of System 1 and System 2. Suppose you are driving to work on the same highway route that you take every day. System 1 is likely to direct your driving actions due to the habitual nature of your regular routine. However, suppose you then remember that you had seen signs last week stating that your usual route would have reduced lanes to accommodate construction. As you consider the several possible alternative routes you might take to get to work, System 2 has taken over directing the decision-making process. The work of System 2 in this example includes both rational and algorithmic processes. That is, you would first have to reflect on your memories to recall the expected construction and then you would have to systematically consider the alternative routes. These two processes have predominantly been considered parts of the same system, thus their operations should correlate highly within individuals. In opposition, studies

have indicated that measures of rational and algorithmic processes do not correlate as well as would be expected if such measures were tapping the same processes (Sá et al., 1999; Stanovich & West, 1997).

Consequently, Stanovich (2012) has developed a tri-partite cognitive model that extends the dual-process account. Figure 2 illustrates the structure of this model. The model further outlines the expected relationships among working memory, other cognitive abilities such as intelligence, and rational thinking dispositions. In contrast to the traditional dual-process model, Stanovich's tri-partite theory divides System 2 into two separate components, while the same single component remains for System 1. Using this structure, Stanovich hypothesizes that the Algorithmic Mind represents individual differences in working memory capability and fluid intelligence, while the Reflective Mind represents people's differences in rational thinking styles and crystallized intelligence. Emphasis is on the Reflective Mind because of its higher level function to initiate an override process (during which the Algorithmic Mind halts responses from the Autonomous Mind). Hence, in the driving example above, the tri-partite model hypothesizes that the operations of System 2 should be divided into two separate types of process. That is, the model predicts that it is actually the Reflective Mind which had recalled the expected occurrence of construction, likely due to the consideration that it is important for you to get to work on time. Subsequently, your Reflective Mind had initiated your Algorithmic Mind to override the routine operations of your Autonomous Mind. Furthermore, the Reflective Mind is hypothesized to initiate processes of simulation in appropriate contexts so that the Algorithmic Mind will perform the decoupling function necessary in order to simulate potential scenarios. Processes of simulation are considered important in all System 2 functioning, partly because contemplating alternative options is critical to knowing when the override function will be effective. In the

driving example, simulation would involve mentally mapping out alternative ways to work. Interestingly, Khemlani and Johnson-Laird (2012) also noted the likelihood that a three part cognitive explanation (control, heuristic, and deliberative) is needed to develop an adequate theory of syllogistic reasoning.

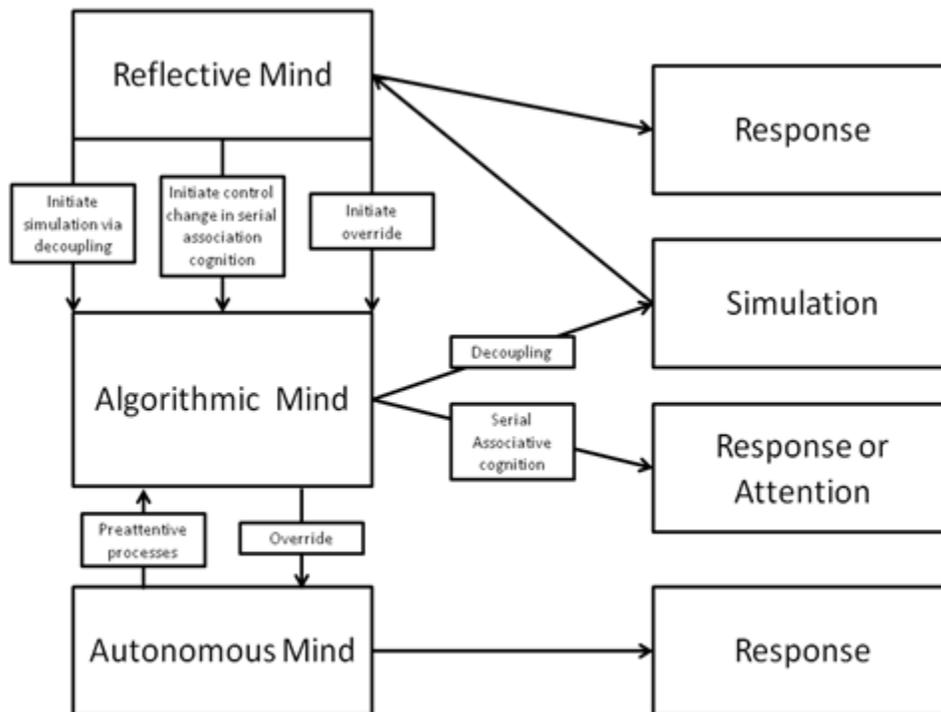


Figure 2. Depiction of Stanovich's (2012) tri-partite cognitive model.

Individual differences have been an important aspect of investigations concerning reasoning abilities. Individual differences in intelligence, working memory, and various thinking styles are correlated with people's reasoning skills and biases in numerous paradigms (Copeland & Radvansky, 2004; Sá et al., 1999; Stanovich & West, 1997). There are some robust effects of individual differences in studies of formal reasoning (Khemlani & Johnson-Laird, 2012). For instance, people with greater intelligence are more able to resist belief-based biases (Stanovich & West, 1997). However, Khemlani and Johnson-Laird (2012) suggested that the connection

between intelligence measures and syllogistic reasoning is not extremely informative because intelligence tests tend to include aspects of similar logical reasoning processes. In addition, people with higher intelligence scores are more likely to have higher working memory capacities (Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005). Still, although working memory measures are correlated with intelligence measures, these constructs are distinct and the former measures exhibit less conceptual overlap with formal reasoning ability compared to the latter measures. Consequently, it is not surprising that working memory is another informative individual difference in regard to explaining syllogistic reasoning ability (Copeland & Radvansky, 2004). On the whole, however, there has been a lack of research focusing on individual differences as predictors of performance in reasoning studies. Hence, requests have continued for further inquiry into such topics (Stanovich & West, 1997; Stanovich, 2012). Stanovich's (2012) tri-partite model predicts that reflective processes precede and manage algorithmic processes, yet each process type is expected to have an independent influence on the relation between source credibility and syllogistic reasoning. In my thesis experiment, working memory served to measure the activity of the Algorithmic Mind, whereas an Actively Open-minded Thinking (AOT) measure and a Cognitive Reflection Test (CRT) quantified the activity of the Reflective Mind.

### **Working Memory and Reasoning Ability**

The working memory construct includes both storage and processing aspects (Baddeley & Hitch, 1974; Baddeley, 2000). This memory system enables information to be held in people's active thoughts so they can further process and manipulate task-relevant information. Working memory is different from short-term memory because the former includes storage and processing elements, whereas the latter only includes storage elements. Baddeley and Hitch's (1974)

original model postulated that the Working Memory system has three components: the *visuospatial sketch pad* (i.e., where visual or spatial information is held), the *phonological loop* (i.e., where phonological or speech-based representations are held) and the *central executive* (i.e., where control of attentional resources is exerted to manage the operations of the other two components).

People's working memory has been measured in diverse reasoning studies, including many paradigms involving syllogisms. For example, Quayle and Ball (2000) investigated working memory in relation to a syllogistic reasoning task which included information content that varied according to believability. In three experiments, these authors assessed working memory using a spatial recall capability measure and an articulatory recall capability measure. These measures were designed to assess Baddeley and Hitch's (1974) visuospatial sketch pad and phonological loop components, respectively. Quayle and Ball predicted that the standard believability by logical validity interaction would depend on what they called "limited working-memory capacity" (p. 1221). They found that belief-bias effects were present for participants with low spatial recall scores, but not for those with high spatial recall scores. They did not find a similar pattern for the articulatory recall measure. However, even if both of their recall measures had been related to belief-bias effects, these measures do not involve both the simultaneous storage and processing aspects of working memory. Because Quayle and Ball did not include a measure of executive functioning (or the central executive in Baddeley and Hitch's 1974 terminology), they seem to be lacking a crucial part of the working memory construct. Hence, it can only be concluded from their results that spatial recall capability limitations are related to greater belief-bias effects. Nevertheless, this finding corroborates the overarching ideas in the theory of mental models. Namely, if people are employing mental representations (and often

relating multiple representations) while attempting to solve syllogisms, it is logical that people's spatial capability would have an important impact on their syllogistic reasoning ability (Johnson-Laird, 1983).

In contrast to the tests of modality-based short-term memory storage just mentioned, working memory span tasks have been found to be excellent measures of executive functioning. Daneman and Carpenter's (1980) complex reading span task was one of the earliest tests of working memory. It involves both processing and storage aspects performed simultaneously. On a trial-by-trial basis, participants must judge if a given sentence is meaningful. After each sentence is presented, participants are given an item (a letter or word) to remember. After two to five sentences have been presented, participants are asked to recall all the items in the order of presentation. Daneman and Carpenter's original work showed that participants' reading span correlated highly with several measures of reading comprehension (including the verbal SAT), whereas the short-term memory measure of word span did not. Hence, these correlations suggested that the reading span was an improvement in accurately measuring working memory.

To ensure that the result did not simply reflect a correlation between the reading component of the reading span task and the reading components of the other reading measures, Turner and Engle (1989) conducted a similar experiment using an operation span task. It was identical to reading span except that participants had to determine if simple mathematical equations are true instead of judging sentences. It revealed that the nature of the secondary task (i.e., judging sentences vs. solving equations) did not impact the correlation of span item recall to reading comprehension, thus consolidating the validity of complex span tasks as measures of working memory. In addition, a counting span task which includes coloured dots as stimuli was developed in order to avoid reliance on language or prior knowledge (Case, Kurland, &

Goldberg, 1982). Using the counting span, working memory span scores could be determined for children and non-native English speakers. Although the reading span has been the most commonly used out of the three, all have been well established as reliable and valid measures of working memory capacity (Kane & Engle, 2002).

The measurement obtained through the complex working memory span tasks is often referred to specifically as Working Memory Capacity (WMC). McCabe, Roediger, McDaniel, Balota, and Hambrick (2010) outlined the similarities between WMC and Executive Functioning (EF). They discussed a common question regarding two major ways of approaching EF: Should the construct of EF be considered unitary or as composed of multiple distinct constructs? McCabe et al. noted that the typical approach to WMC has been to consider it as a unitary construct representing the Central Executive (CE) component of the working memory model. McCabe et al.'s findings indicated that EF and WMC basically measure the same construct, which they refer to as Executive Attention (EA). Moreover, complex working memory span tasks are considered to be excellent measures of EA, whereas EF tasks sometimes include additional task-specific abilities.

Researchers have used working memory span tasks to better understand relations between individual differences in working memory and a variety of higher level cognitive abilities, including reasoning. For instance, Copeland and Radvansky (2004) conducted a study on working memory and syllogistic reasoning in which they used Turner and Engle's (1989) operation span task as their primary working memory measure. Their experimental design involved neutral syllogisms to negate any possible belief-bias effects. Two experiments were conducted in which both working memory and syllogistic reasoning ability were measured. In addition, the relationship between working memory and reasoning ability was compared

according to the number of mental models that syllogisms involved (Johnson-Laird & Bara, 1984). Copeland and Radvansky found that people with larger working memory spans also had better reasoning performance. Moreover, these authors found that the correlation between working memory span and reasoning performance was larger for multiple model syllogisms than for one-model syllogisms. Additionally, response times were found to be negatively correlated with working memory. For one-model syllogisms, people with a larger working memory capacity had faster response times than people with smaller working memory. The correlation was weaker for the multiple-model syllogisms, however. This finding likely indicates that more difficult syllogisms forced participants with smaller working memory resources to use heuristic-based strategies, which reduced their response times.

The aforementioned research establishes the connection among working memory and various aspects of higher cognitive functioning, including syllogistic reasoning ability and the use of mental models during the reasoning process (Copeland & Radvansky, 2004). Furthermore, Copeland et al. (2011) have established a relationship between syllogistic reasoning and the effect of source credibility. Unfortunately, Copeland et al. did not include a measure of working memory. In fact, there does not appear to be an established connection between working memory and source credibility effects in *any* paradigm. Hence, a major goal of my thesis experiment was to investigate the relations between people's working memory and their susceptibility to source expertise biases in a syllogistic reasoning paradigm. Individual differences in working memory provided information regarding how people's algorithmic functioning influences their likelihood to be biased by source expertise. Moreover, this paradigm provided a baseline measurement of the strength of such biases in a strictly structured task that does not encourage people to consider the source expertise information. Finally, the present thesis also expanded such analyses to

investigate the influence of response times in accordance with people's individual differences and source expertise effects.

### **Thinking Dispositions and Reasoning Ability**

Reasoning ability has not only been tied to cognitive abilities such as intelligence and working memory, but it has also been connected to various thinking dispositions. In fact, research has shown that these measures represent relatively distinct categories of individual differences (Stanovich & West, 1997; Stanovich, 2012). Thus, I also included two measures of thinking disposition in my thesis experiment. These measures provided evidence regarding how individual differences in rational thinking processes (i.e., represented by actively open-minded thinking scores and cognitive reflection test scores) relate to source expertise biases in people's reasoning performance.

Dispositional styles of thinking specify the tendencies that shape how people think and perceive information. These styles vary among individuals and predict reasoning performance (Macpherson & Stanovich, 2007; Toplak, West, & Stanovich, 2011). For instance, the Need For Cognition is related to the degree to which people will scrutinize information that they receive (Cacioppo et al., 1983; Petty & Wegener, 1998). Stanovich and West (1997) distinguished among the ways that cognitive capacity (i.e., algorithmic) and thinking dispositions (i.e., rational) predict reasoning processes. They created the Actively Open-minded Thinking (AOT) questionnaire by combining items from several existing conceptually related scales, as well as novel questions the authors devised themselves. The AOT questions measure the strength and durability of people's beliefs and values. This includes how likely they are to consider new and opposing information as valuable. The items were denoted as representing six subscale categories, including: Flexible Thinking (10 questions created by the authors), Openness-Ideas

(eight items from the Revised NEO Personality Inventory; Costa & McCrae, 1992), Openness-Values (eight items from the Revised NEO Personality Inventory; Costa & McCrae, 1992), Absolutism (nine items adapted from the Scale of Intellectual Development; Erwin, 1981; Erwin, 1983), Dogmatism (nine items taken from the short-form of a scale; Rokeach, 1960; Troidahl & Powell, 1965), and Categorical Thinking (three items taken from a subscale; Epstein & Meier, 1989). Scores on the latter three categories were subtracted from the former three in order to represent the concept of an “actively open-minded thinking” style.

Stanovich and West (1997) determined that both AOT scores and cognitive ability scores (the latter being derived from two measures: a person’s SAT score and their score on a vocabulary checklist task) were positively related to performance in a reasoning task (i.e., argument evaluation). Moreover, when submitted to a commonality analysis, each of these measures (i.e., the AOT scores versus the cognitive ability scores) accounted for more unique variance than common variance in the task. Hence, people’s degree of actively open-minded thinking style is shown to have an influence on their reasoning ability distinct from people’s more general cognitive capabilities.

The AOT test was subsequently used in Sá et al. (1999) to assess people’s belief biases on two distinct types of task (syllogistic reasoning and perceptual judgements). With regard to the syllogistic reasoning task, these authors found that both cognitive ability (represented here by WAIS block design and WAIS vocabulary) and AOT significantly predicted belief biases. As participants’ cognitive ability or AOT increased, belief biases decreased (yet to a greater degree for the former than for the latter). Furthermore, Macpherson and Stanovich (2007) investigated the relations between several individual differences and the belief bias in a formal reasoning task as well as the myside bias in two other tasks. Myside bias considers the extent to which people

tend to evaluate hypotheses so that they generate evidence congruent with their own opinions. The results indicated that the Wechsler Abbreviated Scale of Intelligence (WASI), the AOT measure, the Need for Cognition scale, and a Superstitious Thinking scale were all able to predict people's ability to overcome belief biases in a formal reasoning paradigm. That is, people with higher scores (or lower scores in the case of Superstitious Thinking) on these cognitive ability and thinking style measures were more likely to respond in an unbiased manner. However, the individual differences variables failed to predict people's degree of myside bias.

Another important thinking style is measured by the Cognitive Reflection Test (CRT; Frederick, 2005). This thinking style is defined as "...the ability or disposition to resist reporting the response that first comes to mind" (p. 35). That is, the CRT measures a person's tendency to override initial responses and continue reflection to find alternative responses. The three questions included in the measure are similar because each quickly invokes an incorrect, intuitive response which must be ignored or thought through more deeply in order to discover the correct answer (see Table 2 for the items).

Frederick (2005) tested 3,428 participants (across 35 separate studies) in order to verify the reliability and validity of the CRT. Many of these studies also measured other cognitive abilities, such as the Scholastic Achievement Test (SAT), American College Test (ACT), Wonderlic Personnel Test (WPT), and the Need for Cognition (NFC). All had moderate positive correlations with the CRT (ranging from .22 to .46). The moderate strength of these correlations indicates that these tests measure common factors as the CRT, yet they also likely measure distinct characteristics as they were intended. The dependent measures among this set of studies included 17 indices of time preference and 18 indices of risk preference. These questions were partitioned and measured in accordance with four decision-making outcomes: a preference for

the patient option, a preference for the gamble when the expected value favors a gamble, a preference for the gamble when the expected value favors a sure gain, and a preference for the gamble when the expected value favors a sure loss. On the whole, the author found that the CRT was a more consistent predictor of high-quality decision-making than the other four cognitive measures (SAT, ACT, WPT, or NFC). In other words, people higher on the CRT were more patient and typically took more or fewer risks whenever it was beneficial to do so. The fact that the CRT takes much less time to administer to participants than any of the other cognitive measures, and that it also relates more strongly to these major decision-making outcomes, makes this measure attractive for researchers.

Toplak et al. (2011) employed the CRT as a measure of rational thinking style in addition to several thinking disposition measures (including the AOT questionnaire) and several cognitive ability and executive-functioning measures (including a Paced Auditory Serial Addition Test to determine working memory scores). They assessed the relation of these measures to several heuristics-and-biases tasks and to susceptibility to belief-bias in syllogistic reasoning. Their results indicated that high CRT scores were more strongly related to avoiding belief-bias in syllogistic reasoning and to avoiding thinking biases on the heuristics-and-biases tasks than were AOT, working memory, and most of the other measures. Note that the CRT is a performance measure while the AOT questionnaire is a self-report measure; a factor which might account for differences in effectively representing particular thinking styles (Toplak et al., 2011).

Performance measures are more likely to be free from the social desirability concerns that often distort responses to self-report measures. Thus, the authors conclude that the CRT might be a particularly useful measure of people's tendency to engage in rational thought because it is directly assessing the amount of cognitive miserliness in their actions. Nonetheless, the AOT has

still been shown to account for distinctive individual difference information that is not encompassed in various tasks of cognitive capacity or other thinking dispositions.

In accordance with previous research, Stanovich's (2012) tri-partite model predicts that both types of individual differences (i.e., general cognitive ability as represented by working memory capacity and thinking style as represented by AOT and CRT scores) should be related to formal reasoning performance. Moreover, if effects of source expertise occur in a similar manner as other belief-biases, then high levels of both working memory and AOT/CRT scores should be related to low levels of source expertise biases. However, having only a high level of working memory, with low levels of AOT and CRT, is unlikely to correlate with lower levels of source biases because the model leads to the prediction that the reflective system will show its influence before the algorithmic system. That is, due to their association with the early operations of the Reflective Mind, AOT and CRT scores are expected to be more related to the occurrence of source biases in reasoning. Hence, people's recognition that a problem requires their further consideration is a result of them having adequate reflective processes. Accordingly, because working memory is associated with the operations of the Algorithmic Mind, which are only employed as a result of people's realization that a problem requires further consideration, working memory should be related to the responses that people make subsequent to ignoring the source expertise information. Thus, although both high working memory and high AOT/CRT scores might be related to the avoidance of source expertise biases, a person with only high AOT/CRT scores is expected to be more likely to avoid this bias than a person with only high working memory.

In summary, it is clear that source credibility information has the potential to bias people's reasoning. This bias is similar to the belief-bias effect because it stems from people's

prior beliefs (about credibility characteristics). Because participants must reason according to logical validity to be accurate, the syllogistic reasoning paradigm is a useful task to examine the strength of source credibility effects in a situation in which such information is not relevant to their reasoning ability. In addition, there has been very little research linking formal reasoning ability, source credibility and people's individual cognitive differences. Several scholars have explicitly called for more research investigating how people's beliefs about sources affect their reasoning processes (Bromme et al., 2011). Moreover, it is clear that such investigations will be enriched with the inclusion of measures to examine people's cognitive differences.

### **Experiment**

The three major goals of my thesis were as follows. The first goal was to replicate Copeland et al. (2011). I predicted that participants would be less accurate when they evaluated the validity of an invalid syllogism presented by an expert compared to when it was presented by a non-expert.

Secondly, my experiment extended Copeland et al.'s (2011) procedure by presenting source expertise information under two different conditions: a written description only (i.e., the written condition) or a written description presented with an image of the expert (i.e., the visual condition). This manipulation was important for several reasons. Although Copeland et al. found a significant interaction between syllogism type and source expertise with their experimental paradigm, the effect size was small. Furthermore, source information in real-life situations is often provided by other media (e.g., audio, picture, video, or interpersonal interaction) that are generally considered more salient than written information (Wilson & Sherrell, 1993). Hence, increasing the salience of the source information was expected to also increase people's bias for accepting incorrect conclusions given by experts (or rejecting incorrect conclusions given by

non-experts). If the expertise bias is larger with the visual than written medium, then the findings might help explain why many source expertise studies using only written descriptions of expertise have typically yielded small effects.

A third goal of my thesis experiment was to investigate the relations between the source credibility effect and participants' working memory capacity and thinking dispositions. To accomplish this goal, participants were asked to complete a reading span task (Daneman & Carpenter, 1980), the Actively Open-minded Thinking (AOT) test (Stanovich & West, 1997), and the Cognitive Reflection Test (CRT) (Frederick, 2005). I hypothesized that all three measures would be positively correlated with syllogistic reasoning performance overall. Stanovich's (2012) tri-partite model yielded more specific predictions concerning the relations among these measures. According to the tri-partite model, source credibility effects occur when the reflective mind is unable to override the automatic mind. Thus, if people did not realize that their ability to correctly evaluate a syllogism required in-depth consideration, then their responses would likely be faster, automatic, and based on a heuristic such as source credibility. Therefore, fast syllogism evaluation times would indicate that participants are not analysing the logical properties of the problem before providing their answer. Moreover, this pattern of responding should be correlated with low scores on the AOT or CRT more than with low working memory capacity. If people did not realize that it was necessary to analyse a problem, then they would not have used the full extent of their working memory resources to solve it. In contrast, people who were higher on AOT or CRT scores were expected to show longer response times because they would recognize the need to further analyse the syllogisms and to ignore source expertise information.

## Method

### Participants

134 undergraduate students were recruited online through SONA, the Carleton University Psychology Experiment Sign-Up System. This sample size was chosen because it was sufficient to test an effect size of .25 with  $\alpha = .05$  (two-tailed) and power = .95. The power was calculated according to the small effect size that was expected in the written source presentation condition. Participants received 1 percent course credit for completing the experiment.

### Stimuli and Measures

**Syllogistic reasoning task.** Participants evaluated the validity of the syllogisms in one of two conditions: the written condition or the visual condition. The written condition fully replicated Copeland et al.'s (2011) methodology with a few small modifications that are described in this section. It involved the presentation of source descriptions for two characters: an expert and a non-expert. The presentation began with an introductory paragraph establishing a cover story:

You will be presented with reports about groups of people from the small town of River Village. Recently they conducted a survey of hobbies and attitudes (e.g., Do you play chess? What kind of computer do you use?) and all 200 people in the town responded. The first 2 statements in the report are facts that are based on results of the survey. At the end of each report there is a conclusion made by one of two people, Quentin or Zane, that is an interpretation of the facts in the report. Your job is to indicate whether the conclusion made by Quentin or Zane is correct. Before you begin, you will be given descriptions about Quentin and Zane. Please read this information carefully as you will be tested on it right after you read the description. (Copeland et al., 2011, p. 121)

After reading the story, participants viewed the two source descriptions presented in Table 1 in a counterbalanced order. Copeland et al. (2011) created these sources to represent high and low levels of expertise, seemingly with consideration of both the type of education as well as the relevance of the characters' experience to the task of concern in the cover story (i.e., a survey research task). Both characters were conceived, however, to be likeable men of approximately the same age. For the present experiment, the source description for Zane was changed in order to control for characteristics that might be potential confounds to representing the trait of expertise, such as trustworthiness or likeability. The original description of Zane included such characteristics (e.g., it noted that Zane was "friendly", which is not relevant to expertise). However, the description of Quentin remained unaltered.

Table 1  
Source Descriptions Used to Represent High and Low Levels of Expertise

| QUENTIN CARTER   | ZANE ANDERSON   |
|--|---|
| Quentin Carter has been a professor at the local university for 25 years. He has a degree in psychology from Harvard. Quentin conducts survey research on people's attitudes and lifestyles. He has analyzed and published over one hundred scientific papers and has also won numerous research awards. Quentin is well respected by his peers at the university. | Zane Anderson has worked as a mechanic in town for 30 years. He has an Auto Mechanics diploma from the regional college. Zane owns and runs the town's garage. He is really good working with car engines and he is accredited by the Canadian Automobile Association. Zane is well respected by his customers in the area. |

*Note.* These descriptions were adapted from Copeland et al. (2011, pp. 123-124). The description for Zane (the non-expert) was altered in order to control for characteristics that were deemed irrelevant and potential confounds to representing the trait of expertise, whereas the description for Quentin (the expert) remained unchanged.

Following each presentation of a source description, participants were asked to answer three questions about the source. In order to encourage participants to pay attention to the source descriptions, they were told beforehand that these questions would be asked. The questions for the expert source were: "Do Quentin's peers at the university respect him?", "Did Quentin get a psychology degree from Yale?", and "Does Quentin conduct survey research about people's

lifestyles?” The questions for the non-expert source were: “Is Zane really good working with car engines?”, “Do Zane’s customers in the area respect him?”, and “Did Zane get an Industrial Mechanic diploma from college?” Participants pressed either the ‘Y’ key on the keyboard to indicate a “yes” response to each question, or the ‘N’ key to indicate a “no” response.

Presentation order for each set of three questions was randomized.

Once the presentation of the sources was complete, participants proceeded to the syllogistic reasoning task. They were asked to evaluate the validity of 32 syllogisms: 16 necessary syllogisms which were considered “correct” and 16 possible strong syllogisms which were considered “incorrect” (see Evans et al., 1999). Within each category of syllogism, eight had an “a-c” conclusion term order, while the other eight had a “c-a” conclusion term order. Finally, within each set of eight syllogisms, there were two syllogisms from each of the four syllogistic figures. See Appendix A for a complete report of the logical properties of the syllogisms.

Half of the syllogisms were provided by the expert and the other half were presented by the non-expert. For both of these sources, eight syllogisms were necessary and eight were possible strong. In addition, figure and conclusion order were counterbalanced across sources. That is, each set of eight syllogisms (i.e., Expert/Necessary, Expert/Possible Strong, Non-Expert/Necessary, and Non-Expert/Possible Strong), contained two syllogisms written in each figure, and within each figure, both “a-c” and “c-a” conclusion orders were used. Moreover, the pairing of syllogism type with each source was counterbalanced using two sets. The necessary set was divided in half (creating N1 and N2) and the possible strong set was divided in half (creating P1 and P2), which were then combined in two ways (N1/P1 and N2/P2). For any given participant, one of these two sets of syllogisms was paired with one of the sources, whereas the

remaining syllogisms were paired with the other source. That is, if a given participant viewed the N1/P1 syllogisms paired with the expert, then the N2/P2 syllogisms would be paired with the non-expert. Presentation of all syllogisms was fully randomized. Hence, presentation of the two sources followed a mixed design. Finally, note that syllogism difficulty (as measured by the number of mental models needed to evaluate the validity of the conclusions) was equated across stimulus sets (see Copeland et al., 2011; Johnson-Laird & Bara, 1984).

The syllogisms that were used in the experiment were obtained directly from D. E. Copeland (personal communication, August 6, 2012). The authors had originally implemented the syllogisms using terms pertaining to hobbies and status terms. They were careful, however, to create neutral content; that is, conclusions that did not elicit belief-bias effects. This action was necessary in order to avoid a possible confound with the source credibility effect under investigation. Two examples of the syllogisms are given in Figure 3. In addition, in an effort to adapt the syllogisms to Canadian participants, I altered some of the words within eight of them. Only two words were changed: “Obama” became “Harper” and “college” became “university”. This was done in order to maintain the connotations of these expressions (i.e., the political leader of the country that the participants live in, and a post-secondary institution that provides degrees). The full list of syllogisms (with the adjusted content shown) is given in Appendix B.

**Necessary Syllogism**

All cyclists are boaters.  
Some cyclists are wine drinkers.  
Some wine drinkers are boaters.

**Possible Strong Syllogism**

Some boaters are hospital volunteers.  
All chess players are boaters.  
Some hospital volunteers are chess players.

Figure 3. Two examples of the syllogisms used in this experiment.

When the participants began the syllogistic reasoning task, they were told that they would be shown two premises and a conclusion, and that their task would be to determine if the conclusions were logically valid. Participants were reminded of the syllogisms’ source on a trial-

by-trial basis: the source's name and profession were added to each syllogism following the premises and preceding the conclusion (i.e., "(The professor) Quentin's conclusion:" or "(The mechanic) Zane's conclusion:"). The syllogisms were presented one at a time and participants were required to press 'Y' if they thought the syllogism was correct or 'N' if they thought the syllogism was incorrect. No explicit instruction in formal reasoning strategy was provided.

Once all the syllogisms had been evaluated, participants were asked two final questions pertaining to the sources. The first question involved participants' perception of each source's relative accuracy. Participants pressed one of the numerical keys 1 through 7 on the keyboard to rate the accuracy of both sources on a single Likert-type scale (1 = Zane was clearly more accurate, 4 = they were equally accurate, and 7 = Quentin was clearly more accurate). The second question was a memory item used as a manipulation check to ensure that the source information was remembered throughout the reasoning task. Participants were again presented with the source names and asked to indicate which source was the expert. That is, the question read precisely: "Which of the people from this study was considered to be an expert with analyzing surveys, Quentin (the professor) or Zane (the mechanic)?"

Finally, the visual condition was identical to the written condition except for the following modifications. First, when the source descriptions used to represent the expert and non-expert were initially presented, a visual image of the source simultaneously appeared on the computer screen. Similarly, the image of the source appeared beside each syllogism. Hence, each trial included both a written statement denoting the source of the given conclusion as well as the corresponding image of that source.

To obtain the source images for the visual condition, several online databases were searched for photographs that corresponded well with the source descriptions. Two photographs

(one expert and one non-expert) were chosen from iStock (<http://www.istockphoto.com/>). The photos were chosen with the purpose of best representing the source descriptions (i.e., the approximate age, the level of expertise, the specific career, and the amount of experience relevant to survey research) as well as best minimizing extraneous source differences (i.e., how prototypical the image appears in regard to trustworthiness, attractiveness, and likeability). The main difference between the expert and non-expert images was their clothing: the professor is shown in semi-formal attire fitting for academic lectures (e.g., collared shirt, textbook in hand) and the mechanic is shown in a uniform suitable for auto body work (e.g., blue coveralls, wrench in hand). However, face assignment was also counterbalanced across participants in order to control for potential confounds that might result from the different facial characteristics and their combination with the characters' clothing. In this second set of images, the faces were swapped through the use of Adobe Photoshop CS5 software.

In addition, at completion of the experimental tasks described above, participants were asked to rate the images in one of these sets according to five characteristics (i.e., credibility, expertise, trustworthiness, likeableness, and attractiveness). These characteristics had all been shown to influence people's perception of sources in various ways; hence these ratings were obtained in order to account for potential confounds that might have stemmed from unexpected differences between the photos. In the visual condition, participants rated only the image set they had viewed throughout the experiment, whereas in the written condition the image set that was rated was counterbalanced across participants. Participants pressed the keys 1 through 7 in order to make their ratings. These numbers were presented on the screen in accordance with a 7-point scale (1 = strongly unrepresentative, 4 = neither unrepresentative nor representative, and 7 = strongly representative).

**Reading span measure.** The experimenter explained the basic instructions for this test and also told participants to follow all instructions that would appear on the computer screen throughout different parts of the task. The task began with six practice trials identical to the test trials. The test trials consisted of 42 sentence-letter combinations presented individually to participants. The written appearance of each sentence and letter on the screen was accompanied by a simultaneous audio recording. Immediately after reading each sentence and letter, participants were required to judge whether the meaning of the sentence made sense by pressing specified keys on the keyboard. They were also required to memorize the letter that appeared subsequent to each sentence, before proceeding to judge another sentence and memorize another letter. A series of 2 to 5 sentence-letter combinations were presented as a set. Hence, once a total set had been presented, participants were prompted to recall the letters in the same order that they had been presented. Three question marks appeared on the screen to indicate when it was time for participants to write down the letters that they could remember on a paper answer sheet. The number of sentence-letter combinations in any given set was pseudorandomized so participants were unable to anticipate the exact point of recall. The version of reading span that was used in this experiment included sentences taken from Engle (2005). The reading span task was scored by summing all the letters that participants had recalled in the same order as they had been presented (i.e., minimum = 0, maximum = 42). Participants were also scored on their sentence judgment accuracy, with correct responses equal to a score of 1 and incorrect responses equal to 0. Then the number of incorrect responses was summed to determine which participants had not completed the task properly (i.e., those with greater than 6 incorrect responses, which is equivalent to less than 85% accuracy).

**Actively Open-minded Thinking (AOT) measure.** An individual difference measurement of people's thinking style was collected through a self-report questionnaire. Each question in this measure was presented using a Likert-type scale (1 = strongly disagree, 4 = neither agree nor disagree, and 7 = strongly agree). All 41 questions were presented on the computer screen in randomized sequence and participants pressed one of the numerical keys 1 through 7 on the keyboard (representing the Likert-type scale) in order to make each of their responses. A set of particular subscales and items from several questionnaires had been used to form the composite AOT measure (see Appendix C for a full version of the AOT questionnaire). These indexes included questions that were scored normally as well as questions that were reverse-scored. Scores on this measure were calculated in the same manner as in Stanovich and West (1997). That is, after the reverse-scoring was implemented, all responses were summed and averaged to form a single AOT score for each participant. Studies indicate that the 41-item AOT measure has good internal reliability (Cronbach's  $\alpha$  ranges from .81 to .84; Campitelli & Gerrans, 2014; Toplak et al., 2011).

**Cognitive Reflection Test (CRT) measure.** The CRT was used as a second test of thinking disposition. It was taken from Frederick (2005) and consisted of the three questions that are shown in Table 2. The questions appeared on the computer screen consecutively and were presented in randomized order. Participants were required to write their numerical answer to each question on a paper answer sheet. The CRT was scored by summing the number of questions answered correctly (i.e., minimum = 0, maximum = 3). Studies indicate that the 3-item CRT measure has acceptable internal reliability (Cronbach's  $\alpha$  ranges from .60 to .74; Campitelli & Gerrans, 2014; Liberali, Reyna, Furlan, Stein, & Pardo, 2011).

Table 2  
 Questions and Answers for the Cognitive Reflection Test Measure

| Question   | Correct Answer | Intuitive Answer |
|--|----------------|------------------|
| (1) A bat and a ball cost \$1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost? _____ cents   | 5              | 10               |
| (2) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____ minutes   | 5              | 100              |
| (3) In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? _____ days | 47             | 24               |

### Procedure

Participants signed an informed consent before beginning the experiment. Participants were then seated individually in front of a Dell Dimension 5150 desktop computer (with an Intel Pentium 4 processor). PsychoPy software version 1.75.01 (Peirce, 2007; 2008) was used to conduct the first phase of the experiment, whereas E-Prime software version 1.1 (Schneider, Eschman, & Zuccolotto 2002a; 2002b) was used to conduct the second experimental phase.

The first phase of the experiment consisted of three parts. Part one involved the reading span test. Parts two and three included the measures of people's thinking dispositions: the CRT (Frederick, 2005) first and the AOT questionnaire (Stanovich & West, 1997) last. Next, in the second phase, the participants were randomly assigned to the written or visual source expertise condition and then proceeded to complete the syllogistic reasoning portion of the experiment (i.e., cover story instructions, source descriptions with short quizzes, syllogistic reasoning task, two final questions, and the photo ratings). Finally, participants were debriefed, thanked for their participation and granted the course credit. The experiment lasted approximately 1 hour.

## Results

A total of 134 participants completed the experiment. Two participants who did not comply with the syllogistic reasoning task instructions were removed from all analyses. One participant's response times fell below 690 ms in comparison to the sample's average of 20,418 ms. Hence, it is reasonable to infer that he did not read the syllogisms. Another participant answered using the same key for all problems except three, making this participant the only outlier on accuracy in the sample. Hence, the final sample included 132 participants (47 male, 85 female), with an approximate mean age of 22 ( $SD = 7.83$ , range = 17 to 64).

### Syllogistic reasoning task

The internal consistency among all 32 syllogisms in the reasoning task yielded a Cronbach's  $\alpha$  of .63. This value indicates that the 32 items were shy of having good reliability (Kline, 2000). Yet, in the context of the present experimental replication, this value is acceptable. It is also of note that Cronbach's  $\alpha$  differed in accordance with the syllogism type (i.e., the value was .61 for the necessary syllogisms and .70 for the possible strong syllogisms). This difference in reliability between the syllogism types is likely explained by the different numbers of mental models associated with each type. Specifically, the necessary syllogisms involved either one or three mental models to be considered (making some of these syllogisms clearly more difficult to solve than the others), whereas all except one of the possible strong syllogisms involved two mental models.

The primary analysis for this experiment evaluated whether source expertise effects would vary by the written or visual mediums of presenting the source information. A  $2 \times 2 \times 2$  (Syllogism Type [necessary, possible strong]  $\times$  Source [expert, non-expert]  $\times$  Medium [written, visual]) mixed-design analysis of variance (ANOVA) was conducted with accuracy (i.e.,

proportion correct) as the dependent variable. The within-subjects factors included syllogism type and source, whereas medium was a between-subjects factor. The results are shown in Figure 4. The findings of this analysis revealed a significant main effect of syllogism type,  $F(1, 130) = 280.99$ ,  $MSE = .07$ ,  $p < .001$ ,  $\eta_p^2 = .68$ . That is, in accordance with an extensive prior literature, participants were more accurate at solving the necessary syllogisms than they were at solving the possible strong syllogisms. In addition, there was a marginally significant Syllogism Type  $\times$  Medium interaction,  $F(1, 130) = 3.59$ ,  $MSE = .07$ ,  $p = .06$ ,  $\eta_p^2 = .03$ . Simple main effects indicated that participants were marginally more accurate in the written condition compared to the visual condition, but only among the necessary syllogisms,  $F(1, 130) = 3.07$ ,  $MSE = .03$ ,  $p = .082$ ,  $\eta_p^2 = .02$ . No other effects were significant (all  $ps > .05$ ). Hence, neither media led to a replication of Copeland et al.'s (2011) key finding: participants were not more likely to accept the conclusions of possible strong syllogisms when they were presented by an expert. Nonetheless, it should be noted that 16 participants (12.1%) did not respond accurately to the manipulation check question at the end of the syllogistic reasoning task (i.e., asking them to identify which source was the expert). Hence, it was possible that removing these participants from the analyses would alter the results. This was not the case, however.

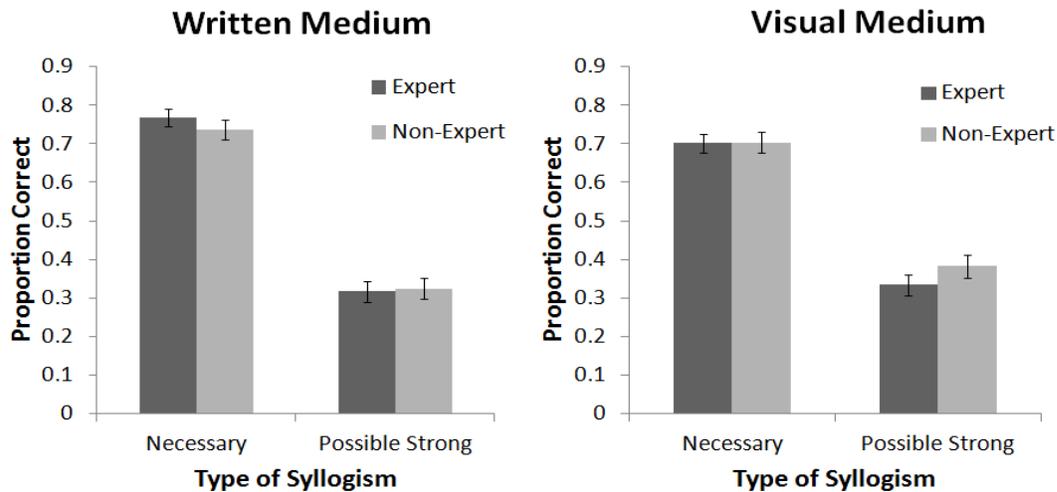


Figure 4. Proportion correct for response accuracy across the four combinations of syllogism type and source, shown separately for each medium of presentation. Error bars represent standard errors.

The  $2 \times 2 \times 2$  mixed-design ANOVA was repeated with response time (RT in milliseconds) as the dependent variable. The results are shown in Figure 5. There was a main effect of syllogism type,  $F(1, 130) = 17.71$ ,  $MSE = 27580000$ ,  $p < .001$ ,  $\eta_p^2 = .12$ . That is, participants had longer RTs for possible strong syllogisms than for necessary syllogisms. No other effects reached significance (all  $ps > .05$ ).

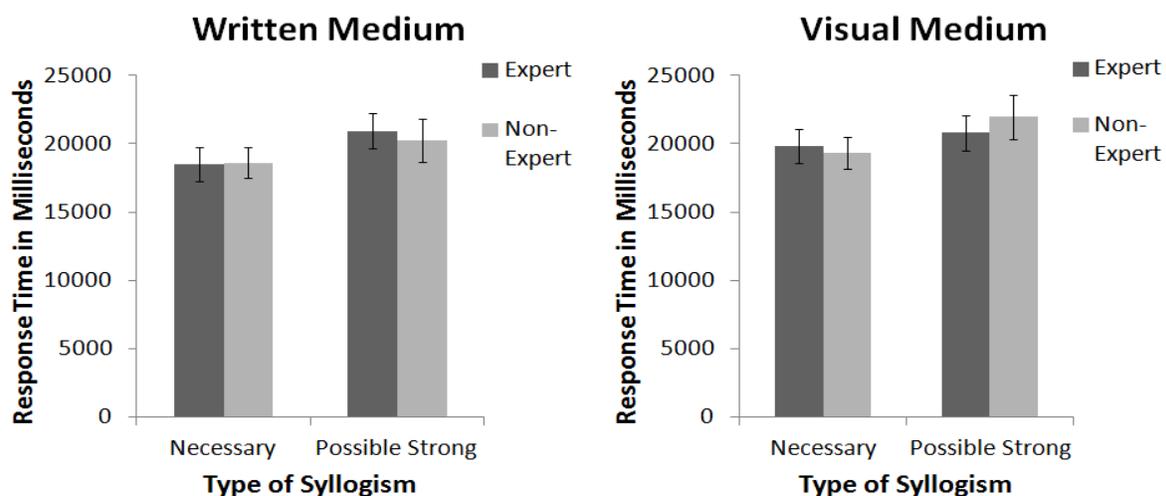


Figure 5. Response time in milliseconds across the four combinations of syllogism type and source, shown separately for each medium of presentation. Error bars represent standard errors.

Finally, recall that participants had been asked to rate their perception of each source's relative accuracy on a 7-point Likert-type scale. A one sample t-test indicated that the rating ( $M = 3.92$ ,  $SE = .11$ ) was not significantly different from "4",  $t(131) = -.68$ ,  $p = .496$ . Hence, Quentin and Zane were considered to be equally accurate. This result did not replicate the significant difference in perceived accuracy that Copeland et al. (2011) found, which had indicated that their participants thought the expert was more accurate than the non-expert. However, the unbiased perceived accuracy rating in the present experiment is consistent with the fact that no source expertise bias was found among the full sample of participants.

### **Relationships among the individual differences measures and source expertise effects**

The second set of analyses investigated the relations among the individual differences measures (i.e., WM scores, AOT scores, and CRT scores) and any potential source expertise effects. To conduct such analyses, a mean source expertise bias score was first calculated for each participant using the following formula:

$$\text{Source expertise bias} = (\text{Necessary/Expert} - \text{Necessary/Non-Expert}) + (\text{Possible Strong/Non-Expert} - \text{Possible Strong/Expert})$$

Note that for the necessary syllogisms, participants' accuracy for the non-expert syllogisms was subtracted from their accuracy for the expert syllogisms. In contrast, for the possible strong syllogisms, participants' accuracy for the expert syllogisms was subtracted from their accuracy for the non-expert syllogisms. The difference in these calculations is due to the expected difference in accuracy between necessary versus possible strong syllogisms for a participant who shows the source expertise bias. That is, participants showing the source expertise bias should be more inclined to accept as valid syllogisms provided by the expert regardless of whether such syllogisms are actually valid. By doing so, these participants should

be more accurate when the syllogisms are indeed valid (i.e., necessary), but less accurate when the syllogisms are invalid (i.e., possible strong). Hence, participants who show the largest differences in the expected directions are those who displayed the highest source expertise bias. Finally, note that the data from the written and visual conditions were merged for these analyses.

One-tailed Pearson product-moment correlations were conducted to analyze relations among the individual differences measures and the mean bias score. In addition, correlations were conducted with participants' mean accuracy (i.e., the average accuracy for all syllogisms) and mean RT (i.e., the average RT for all syllogisms). One-tailed correlations were chosen rather than two-tailed because specific correlations were predicted. More specifically, the three individual differences measures were expected to be positively related to mean accuracy and mean RT, and yet be negatively related to mean bias. Mean accuracy was also predicted to be positively related to mean RT. The results for these correlational analyses as well as the descriptive statistics for all the measures are shown in Table 3. It may first be observed that AOT scores, CRT scores and WM scores were not significantly related to each other (all  $ps > .05$ ). Hence, these measures might best be considered to represent three distinct constructs or separable aspects of cognitive functioning. As expected, mean accuracy and mean RT were positively correlated,  $r(132) = .46, p < .001$ . Moreover, in accordance with predictions, the CRT scores were positively correlated with mean accuracy,  $r(132) = .22, p = .005$ , as well as with mean RT,  $r(132) = .17, p = .027$ . In contrast with expectations, however, AOT scores were positively correlated with mean bias,  $r(132) = .19, p = .014$ . No other correlations were significant (all  $ps > .05$ ). Finally, note that the AOT measure was found to have good reliability (Cronbach's  $\alpha = .84$ ).

Table 3  
Correlations among Individual Difference Measures and Mean Outcome Scores

| Measure          | Mean  | SD    | 1    | 2     | 3    | 4   | 5     |
|------------------|-------|-------|------|-------|------|-----|-------|
| 1. AOT           | 4.94  | .59   | -    |       |      |     |       |
| 2. CRT           | .25   | .36   | .11  | -     |      |     |       |
| 3. WM            | 35.56 | 5.75  | -.08 | .08   | -    |     |       |
| 4. Mean Bias     | .04   | .33   | .19* | -.14  | .01  | -   |       |
| 5. Mean Accuracy | .53   | .13   | .07  | .22** | -.01 | .03 | -     |
| 6. Mean RT       | 20418 | 11006 | .02  | .17*  | .07  | .00 | .46** |

*Note.* Aside from the mean and SD values, all of the values shown are one-tailed Pearson product-moment correlation coefficients ( $r$ ). AOT = Actively Open-minded Thinking; CRT = Cognitive Reflection Test; WM = Working Memory.

\* $p < .05$ ;

\*\* $p < .01$

Although the correlation between participants' WM and the various outcome scores was not significant, the previous correlations did not account for the sentence evaluation aspect of the WM task. Twelve participants (9.1%) did not maintain the minimum acceptable criterion of 85% accuracy in the sentence evaluation aspect of the task (see Conway et al., 2005) and thus their working memory capacity scores were likely invalid. Thus, the correlations were re-run after removing these participants ( $n = 120$ ). The results did not change, however.

Participants who had inadequate accuracy on the sentence evaluation aspect of the WM task and those who did not answer the manipulation check question correctly were combined into a more general category of participants who did not follow the experimental instructions properly. Then the  $2 \times 2 \times 2$  (Syllogism Type [necessary, possible strong]  $\times$  Source [expert, non-expert]  $\times$  Medium [written, visual]) mixed-design ANOVA on accuracy was re-run with this sub-sample of participants ( $n = 106$ ). The main effect of syllogism type,  $F(1, 104) = 272.42$ ,  $MSE = .07$ ,  $p < .001$ ,  $\eta_p^2 = .72$ , and the marginally significant Syllogism Type  $\times$  Medium interaction,  $F(1, 104) = 3.14$ ,  $MSE = .07$ ,  $p = .079$ ,  $\eta_p^2 = .03$ , remained similar to the full sample analysis. In contrast to the full sample, however, the Syllogism Type  $\times$  Source interaction became marginally significant,  $F(1, 104) = 3.69$ ,  $MSE = .03$ ,  $p = .057$ ,  $\eta_p^2 = .03$ . In addition, the

Source  $\times$  Medium interaction became significant,  $F(1, 104) = 4.38$ ,  $MSE = .02$ ,  $p = .039$ ,  $\eta_p^2 = .04$ . No other effects reached significance (all  $ps > .05$ ). Thus, because the removal of these participants resulted in the marginal interaction between source and syllogism type, which represents the source expertise effect, it is likely that at least a portion of these participants had not paid attention to the instructions during the reasoning task or the experiment overall. That is, their lack of attention might explain the lack of effect in the full sample. However, it is difficult to determine which participants did not pay attention to the task instructions in general and which ones simply did not pay attention to the source information while solving the syllogisms. The latter option is an alternative that could have resulted in participants' forgetting the source descriptions and thus also could explain their lack of source expertise effect. Nevertheless, because there were 28 participants in this category (i.e., 21.2%), it did not seem appropriate to exclude them all due to the expectation that some of them likely completed the experiment improperly. If a portion of these participants actually did complete the experiment properly, then this action might constitute a distortion of the dataset. Hence, to err on the cautious side, the full sample was maintained for all further analyses.

To further understand the significant correlations involving the two thinking style measures, several additional exploratory analyses were performed. First, a regression was conducted in order to determine whether the three individual difference variables together (i.e., CRT, AOT, and WM) would predict mean bias scores better than each had individually. Results indicated that the three predictor model explained a significant proportion of variance in mean bias scores,  $R^2 = .06$ ,  $F(3, 128) = 2.92$ ,  $p = .037$ . AOT scores were significantly positively associated with mean bias,  $\beta = .21$ ,  $t(128) = 2.46$ ,  $p = .015$ . CRT scores marginally significantly predicted mean bias,  $\beta = -.17$ ,  $t(128) = -1.92$ ,  $p = .057$ . WM scores did not significantly predict

mean bias,  $\beta = .04$ ,  $t(128) = .43$ ,  $p = .667$ . Interestingly, the effects of the thinking style measures were in opposite directions. In accordance with predictions, as CRT scores increased mean bias decreased. Yet, unexpectedly, as AOT scores increased mean bias also increased.

Two more regressions were run with the criterion variables of mean accuracy and mean RT. Again, the three individual differences were entered as predictors. Results indicated that the three predictor model only explained a marginally significant proportion of variance in mean accuracy scores,  $R^2 = .05$ ,  $F(3, 128) = 2.38$ ,  $p = .073$ . The CRT was the only measure significantly positively associated with mean accuracy,  $\beta = .22$ ,  $t(128) = 2.54$ ,  $p = .012$ . For mean RT scores, results indicated that the three predictor model did not explain a significant proportion of variance,  $R^2 = .03$ ,  $F(3, 128) = 1.41$ ,  $p = .243$ . However, the CRT marginally significantly predicted mean RT,  $\beta = .16$ ,  $t(128) = 1.86$ ,  $p = .065$ . Although small, both of the trends involving the CRT were in the expected direction. That is, people with higher CRT scores were more accurate overall and they also took somewhat longer to respond to the syllogisms.

Although mean RT and mean bias were already analyzed separately in regard to the three individual difference scores, another important endeavour was to compare participants' RTs to the relations between their individual differences and any potential source expertise effects. Moreover, there was a prior interest in comparing participants who had slower RTs to those who had faster RTs. To do this, a median split was conducted on the mean RT variable in order to separate the participants into two groups (fast vs. slow RT). Then, because prior predictions had been made, one-tailed correlational analyses were performed separately for each group. Firstly, within the slow RT group, mean RT was positively correlated with mean accuracy,  $r(66) = .42$ ,  $p < .001$ . Also within this group the results showed that CRT scores were positively correlated with mean accuracy,  $r(66) = .31$ ,  $p = .006$ . These results were in accordance with the

experimental hypotheses. However, in contrast to experimental hypotheses, also within the slow RT group AOT scores were positively correlated with mean bias,  $r(66) = .26, p = .017$ . Note that these relations were similar to the results found among the full sample.

On the other hand, within the fast RT group the CRT was found to be negatively correlated with mean bias,  $r(66) = -.27, p = .014$ . Also within this group mean RT was positively correlated with mean accuracy,  $r(66) = .22, p = .04$ . No other correlations for either the slow or fast RT groups reached significance (all  $ps > .05$ ). Although the negative correlation between CRT and mean bias (within the fast RT group only) had not been found in the full sample, there was a marginal trend showing this relationship in the results from the regression on mean bias. Hence, dividing participants according to their mean RT made it clear that the trend between CRT scores and mean bias exists, but is only apparent among participants with fast RTs. Moreover, this relationship is compatible with one of the experimental hypotheses which stated that participants who had both low CRT scores and fast RTs would be more likely to exhibit the source expertise bias.

#### **Additional analysis of the CRT in relation to the source expertise effect**

Because the CRT was related to all the mean dependent measures in the predicted directions, I further investigated the connection between this individual difference and the source expertise bias. Participants were divided into high and low CRT score groups. Specifically, participants who answered 2 or 3 questions correctly were placed in the high CRT group ( $n = 33$ ), whereas participants who answered 0 or 1 question(s) correctly were placed in the low CRT group ( $n = 99$ ).

The  $2 \times 2 \times 2$  (Syllogism Type [necessary, possible strong]  $\times$  Source [expert, non-expert]  $\times$  Medium [written, visual]) mixed-design ANOVA on accuracy was repeated separately for the

high CRT versus the low CRT groups. Results for the high CRT group indicated only a significant main effect of syllogism type,  $F(1, 31) = 75.26$ ,  $MSE = .06$ ,  $p < .001$ ,  $\eta_p^2 = .71$ . Results for the low CRT group, however, indicated the significant main effect of syllogism type,  $F(1, 97) = 190.91$ ,  $MSE = .08$ ,  $p < .001$ ,  $\eta_p^2 = .66$ , as well as a significant Syllogism Type  $\times$  Source interaction,  $F(1, 97) = 6.35$ ,  $MSE = .03$ ,  $p = .013$ ,  $\eta_p^2 = .06$ . This interaction shows the same pattern of a source expertise bias that Copeland et al. (2011) found. Results from these separate analyses are shown in Figure 6. Also note that among both the high and low CRT groups, the results did not indicate the marginal Syllogism Type  $\times$  Medium interaction that the full analysis had. No other effects were significant (all  $ps > .05$ ).

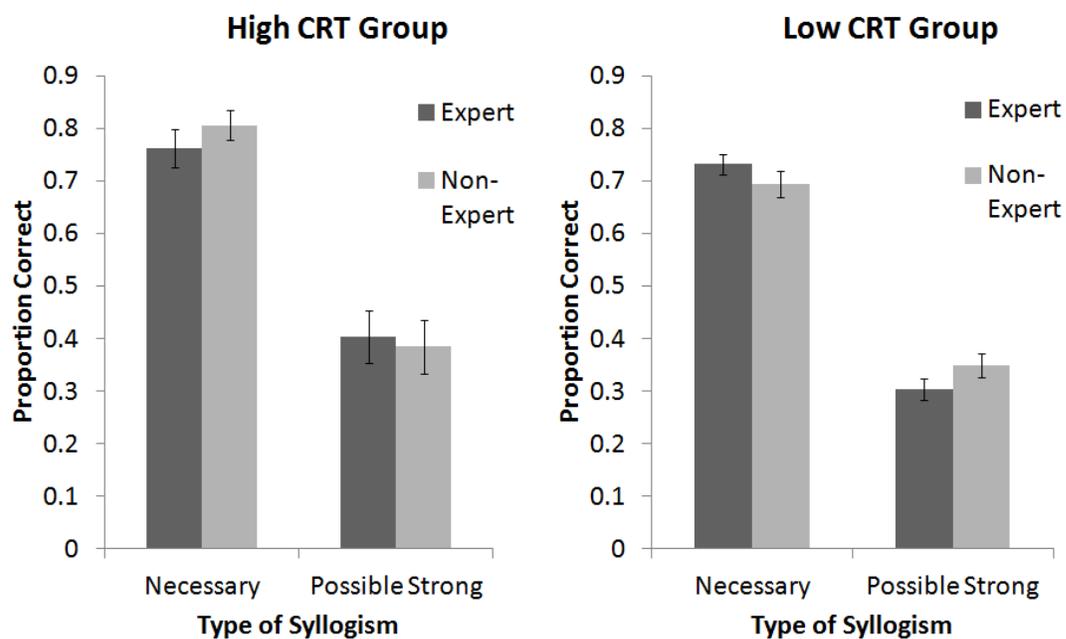


Figure 6. Results showing an absence of the source expertise effect in the high CRT group (left graph) compared to the presence of this effect in the low CRT group (right graph). Error bars represent standard errors.

### Participants' evaluations of the pictures depicting the expert and non-expert

One last series of analyses were conducted to investigate the photo ratings that all participants gave at the end of the experiment. Participants in the visual condition rated the

photos that they had seen throughout the entire reasoning task (i.e., half rated the original photos and half rated the altered photos). Those in the written condition were randomly assigned to rate either the original or the altered photos. Each photo was rated on five traits: credibility, trustworthiness, expertise, likeableness, and attractiveness (see Table 4 for descriptive statistics).

Table 4  
Mean Ratings of the Original and Altered Expert and Non-Expert Photos on Five Traits

| Trait       | Original    |             | Altered     |             |
|-------------|-------------|-------------|-------------|-------------|
|             | Expert      | Non-Expert  | Expert      | Non-Expert  |
| Attractive  |             |             |             |             |
| Written     | 2.94 (1.35) | 2.52 (1.23) | 2.35 (1.13) | 2.85 (1.44) |
| Visual      | 2.91 (1.83) | 2.85 (1.56) | 2.47 (1.44) | 2.56 (1.52) |
| Expert      |             |             |             |             |
| Written     | 5.12 (1.41) | 4.76 (1.48) | 5.41 (1.08) | 4.97 (1.22) |
| Visual      | 5.88 (1.36) | 5.00 (1.60) | 5.34 (1.68) | 4.81 (1.87) |
| Trustworthy |             |             |             |             |
| Written     | 4.91 (1.36) | 4.85 (1.15) | 4.71 (1.14) | 4.50 (1.31) |
| Visual      | 5.09 (1.67) | 5.64 (1.32) | 5.19 (1.28) | 5.31 (1.20) |
| Likeable    |             |             |             |             |
| Written     | 5.00 (1.12) | 5.15 (1.58) | 5.41 (0.96) | 5.12 (1.37) |
| Visual      | 5.03 (1.45) | 6.03 (1.33) | 5.31 (1.09) | 5.53 (1.11) |
| Credible    |             |             |             |             |
| Written     | 5.42 (1.06) | 4.52 (1.03) | 5.35 (1.13) | 4.65 (0.98) |
| Visual      | 5.91 (1.13) | 5.33 (1.38) | 5.81 (0.97) | 4.97 (1.64) |

*Note.* Values in the cells indicate the mean ratings (and standard deviations) for each photo and trait combination separated according to the medium of source presentation. The ratings were made on a 7-point scale (1 = strongly unrepresentative, 4 = neither unrepresentative nor representative, and 7 = strongly representative).

Ratings on these traits constituted multiple dependent variables. A  $2 \times 4$  (Medium [visual, written]  $\times$  Photo [original non-expert, original expert, altered non-expert, altered expert]) multivariate analysis of variance (MANOVA) was performed. Results showed a main effect of medium, Wilks' lambda = .93,  $F(5, 252) = 3.73$ ,  $p = .003$ ,  $\eta_p^2 = .07$ , and a main effect of photo, Wilks' lambda = .82,  $F(15, 762) = 3.28$ ,  $p < .001$ ,  $\eta_p^2 = .06$ . Regarding the effect of medium, participants in the visual condition had significantly higher ratings on trustworthiness ( $p = .001$ ) and credibility ( $p < .001$ ), and marginally significantly higher on likeableness ( $p = .051$ ), than

participants in the written condition. These effects indicated that the photos had an overall greater appeal to participants in the visual condition, which is not surprising considering that these participants had viewed the photos on every trial beside the syllogisms. Thus, these participants were more familiar with the photos than were participants in the written condition which makes them more likely to develop preferences for them (i.e., according to the classic psychological phenomenon known as the “mere-exposure effect”; Zajonc, 1968). Hence, this phenomenon might explain their tendency to give higher ratings.

Regarding the effect of photo, there were several differences among the four photos. Firstly, the original expert photo was rated as having more expertise than both the original non-expert photo ( $p = .016$ ) and the altered non-expert photo ( $p = .019$ ). In addition, the altered expert photo was rated as having marginally more expertise than the original non-expert photo ( $p = .053$ ) and the altered non-expert photo ( $p = .06$ ). Furthermore, both the original and altered expert photos were rated as having more credibility than the original and altered non-expert photos (all  $ps < .002$ ). All of these differences were in the expected direction, indicating that the photos were adequate representations in regard to the desired trait of expertise (and more generally, the desired trait of credibility). Additionally, the original expert photo was rated as more attractive than the altered expert photo ( $p = .043$ ), and the original non-expert photo was rated as more likeable than the original expert photo ( $p = .01$ ). The difference in attractiveness between the original expert versus the altered expert was not of particular alarm because the major concerns were only in regard to comparisons between the expert photos and the non-expert photos. Hence, the difference in likeableness between the original non-expert and the original expert was undesirable. Thus, several independent samples t-tests were conducted to determine whether accuracy differed between the original versus the altered sets of photos (i.e., for the

visual condition participants only;  $n = 65$ ), yet none of the differences were significant for either syllogism type provided by the non-expert or the expert (all  $ps > .05$ ).

### **General discussion**

The goal of the present thesis was to explore the situational and individual difference variables that influence people's degree of source expertise bias in a syllogistic reasoning paradigm. The experiment was designed to replicate Copeland et al.'s (2011) basic methodology. Nevertheless, it also extended it by manipulating the medium of the source that presented the information and by measuring people's working memory and thinking dispositions. The results with the full sample of participants indicated that there was no source expertise effect (i.e., no tendency to accept the expert-provided syllogisms or to reject the non-expert syllogisms). Hence, Copeland et al.'s results were not replicated. Furthermore, visual and written source media did not yield a difference in the source expertise bias. However, the Cognitive Reflection Test (CRT) did seem to explain why Copeland et al.'s results were not replicated: participants with low scores on the CRT exhibited the source expertise bias, whereas those with high scores did not. This result was compatible with expectations and indicates that this individual difference measure is critical in explaining the tendency for people to rely on heuristic source expertise characteristics versus to exert additional cognitive effort to reason carefully. Hence, the relationship between this thinking style and people's tendency to be biased by source expertise while reasoning with syllogisms is explored fully in this discussion. In a contrasting result, people with high scores on the Actively Open-minded Thinking (AOT) measure displayed the source expertise bias, whereas people with low scores did not. This finding was unexpected and is discussed in more detail later. Finally, Working Memory (WM) was unrelated to the source expertise bias.

There are several potential reasons that might explain why the present experiment's results did not replicate Copeland et al. (2011). Firstly, it is possible that the source expertise

effect does not exist among the general population. Hence, Copeland et al.'s results might simply be due to chance. However, because source credibility effects have been found in a wide array of studies in the literature, this explanation is unlikely. More likely is the possibility that the source credibility effect as it occurs in a syllogistic reasoning paradigm is small. Hence, it will sometimes be difficult to find. I return to this explanation later.

The second reason that the present experiment might have failed to replicate Copeland et al. (2011) is methodological. Some changes were made. For instance, a few syllogisms were slightly modified to adapt them to Canadian participants. More importantly, recall that one of the source descriptions (Zane, the non-expert) was altered slightly. Specifically, because the Zane description seemed to include information about characteristics extraneous to the intended trait of expertise, a few sentences in the Zane description were changed to focus more closely on expertise. Although it is unlikely that these small changes represented important methodological differences, it is still a possibility that they affected participants' perceptions of the sources enough to alter their responding so that they became less biased by the source expertise information.

A third explanation is that the source expertise effect only occurs among a sub-sample of the population. If this is the case, it would explain why the effect has been small and difficult to obtain. It would also raise the question: what characteristics might differentiate people who show the effect versus those who do not? Luckily, the inclusion of individual differences measures in the present experiment has proven useful in shedding light on this query. In particular, the results involving the relationship between the CRT and the source expertise effect were especially informative about the nature of the bias in syllogistic reasoning. As has been already noted, the thinking disposition represented by the CRT involves people's tendency to inhibit intuitive

responses and to search for alternative solutions to a problem. Hence, the results of the present experiment indicate that this thinking style is clearly related to the specific likelihood that people will ignore irrelevant source information and exert the mental effort required to solve syllogisms.

Overall, the CRT scores were related to people's accuracy, response times, and their tendency to exhibit the source expertise bias. However, the source expertise bias was only found among people who scored 0 or 1 on the CRT (i.e., the low CRT group). Those who scored 2 or 3 (i.e., the high CRT group) did not exhibit the bias. Nevertheless, the sample size in the high CRT group was fairly small overall. Thus further study with a larger sample size of participants who score highly on the CRT will be required to determine whether the lack of source expertise bias remains consistent in this group of individuals. Furthermore, additional analysis on response times indicated that only people with fast response times showed a relationship between CRT score and source expertise bias. That is, people who responded quickly to the syllogisms were more likely to show the bias if their CRT score was low. This result is not surprising because it was expected that people who were susceptible to the source expertise bias would both respond to the syllogisms quickly and have low CRT scores. This is because when people rely on easily available heuristic cues to help them make responses, they are likely to respond more rapidly, are less likely to search for alternative solutions, and are often more biased in their judgements.

The unexpected relationship between participants' AOT scores and their degree of source expertise bias should be interpreted cautiously. The prior literature has shown that people with higher AOT scores are less susceptible to belief-biases in syllogistic reasoning (Sá et al., 1999; Macpherson & Stanovich, 2007; Toplak et al., 2011). In contrast, the present results indicated that people with higher AOT scores are *more* susceptible to the source expertise bias. Although this finding is inconsistent with the literature, the correlations among individual differences

measures in the present experiment were relatively consistent with Toplak et al.'s (2011) results. That is, the AOT measure and the CRT had a correlation of .10 in Toplak et al. (2011), which is similar to the correlation of .11 found in the present experiment. In addition, although different tasks were used to represent WM across the two studies, the correlation between WM scores and the AOT measure was .02 in Toplak et al. (2011) compared to -.08 in the present experiment. Thus, because of the consistency of previous findings between the AOT scale and bias in reasoning, it seems likely that the unexpected relationship between AOT scores and the source expertise bias in the present experiment simply occurred due to chance. Hence, it might not prove replicable in future experiments.

On the other hand, it should be remembered that the bias score assessed in the present experiment was in regard to the influence of the source information, rather than the content of the syllogisms as is the case with belief-bias effects. Hence, if the positive correlation between AOT and the source bias is replicable, then it should be considered that the source credibility effect has a distinct relationship to AOT in comparison to the relationship between AOT and belief-bias effects. That is, it is possible that people's AOT scores are related to a tendency to think about source information in a divergent manner compared to how they think about the information within syllogisms. For instance, although high AOT scorers tend to reason more appropriately with syllogistic content than do low AOT scorers, perhaps they also have a greater tendency to believe that source information is relevant in most situations. Thus, an occurrence of both tendencies would account for the fact that people who score high on AOT might be less likely to be biased by syllogistic content yet more likely to be biased by source expertise information.

Additionally, if the unexpected result is replicable, it is possible that people's degree of social desirability might have influenced their responses on the AOT scale. That is, people who scored higher on the AOT scale in the present experiment might also be those with larger social desirability biases. This explanation is a possibility in particular because of the structure of the procedure in the present experiment. Specifically, the AOT measure was administered to all participants after they had already completed the WM measure and the CRT. Because these two performance measures likely require a moderate degree of cognitive effort and many participants find the WM task to be difficult, it is possible that completing these tasks caused participants to become concerned about the strength of their overall cognitive abilities. Thus, once they were finally able to provide simple self-report indications of their abilities within the AOT measure, they might have attempted to boost the appearance of their cognitive skills through responding in a socially desirable manner. Therefore, the inclusion of a social desirability scale in experiments using the AOT measure is a potential direction for future testing to explore. However, finding a relationship between higher AOT and greater social desirability might not be as compatible with the findings that higher AOT leads to less belief-bias. Thus, an explanation based on the influence of social desirability is not likely to be fully adequate in accounting for the present result as well as the prior literature.

Likewise, the results involving WM scores in relation to reasoning ability in the present experiment were not consistent with prior literature. WM scores were not found to be correlated to mean accuracy in the syllogistic reasoning task. This result is inconsistent with Copeland and Radvansky's (2004) finding that people with higher scores on an operation span task also had higher accuracy at solving syllogisms. In addition, previous research found that WM scores (as measured by a Paced Auditory Serial Addition Test) were correlated with CRT scores as well as

with the ability to avoid belief-bias in syllogistic reasoning (Toplak et al., 2011), but the present study did not obtain these relationships. These discrepancies indicate that there might be an issue with the WM results in the present experiment. A specific aspect of the present methodology might have altered the typical relations between WM and syllogistic reasoning performance, such as the inclusion of additional information (i.e., about the sources) on each trial. Copeland and Radvansky designed their experiment to isolate the correlation between WM and syllogistic reasoning performance, whereas maybe the additional variables in the present experiment concealed this relationship. Thus it is clear that further experimentation will be required to enable an interpretation of these results.

In regard to the limitations of my thesis, there are several potential concerns with the results of the present experiment. First of all, there is the problem of explaining the marginal interaction between the medium of presentation and the type of syllogism. Recall that accuracy on the necessary syllogisms was marginally higher for people who viewed only written source information compared to people who also viewed visual source information. One possibility for the difference is that the additional visual information might simply have been more of a distraction to participants than only the written information. The increased distraction might have caused the decreased accuracy because distraction has been shown to disrupt performance (Kiesler & Mathog, 1968; Petty, Wells, & Brock, 1976). However, after consideration of the CRT scores, another likely explanation for this finding became clear. Unfortunately, the random assignment of participants into visual versus written conditions did not create an equal distribution of high and low CRT scorers (i.e., those who scored 0 or 1 versus those who scored 2 or 3). Although the random process placed a relatively adequate ratio of low CRT scorers into both mediums (i.e., the written condition  $n = 44$  and the visual condition  $n = 55$ ), there was a

large discrepancy in the number of high CRT scorers placed into the visual versus written conditions ( $n = 10$  versus  $n = 23$ , respectively). Hence, this discrepancy is a plausible explanation for why there was the marginal trend for people in the written condition to have higher accuracy than people in the visual condition. Because CRT scores were positively correlated with mean accuracy, and because more people in the written condition were high CRT scorers, the written condition would have somewhat higher accuracy than the visual condition. The fact that the higher accuracy was only found among the necessary syllogisms is another question, but it does not discount the explanation just outlined. In addition, when the analyses were run separately for high versus low CRT groups, the marginal interaction between syllogism type and medium of presentation disappeared for both groups. This supports the contention that the difference in CRT scores between the visual and written conditions is the most sensible explanation for the interaction finding. Hence, if a future study matched the CRT scores evenly across medium of presentation, it is unlikely that there would be a difference in accuracy between media.

Another issue in the present experiment involves the results of the manipulation check question. Among the 132 participants, there were 16 participants who answered this simple question incorrectly (i.e., 6 from the written condition and 10 from the visual condition). This is a large number compared to the one person who answered the question incorrectly in Copeland et al.'s (2011) sample of 60 participants. One explanation of this finding is that these participants did not remember which character was the expert, which implies that they did not read the source descriptions carefully, nor did they pay attention to the source information throughout the whole reasoning task. Hence, the fact that these participants did not attend to the instructions and source

information constitutes an obvious reason why they would not have exhibited the source expertise bias.

In addition, recall that 12 participants did not follow the instructions for the WM task properly. Moreover, recall that participants who did not answer the manipulation check question correctly and those who had inadequate accuracy on the sentence evaluation aspect of the WM task were combined into a more general category of participants who did not follow the experimental instructions properly. Once all 28 of these participants were removed and the mixed-design ANOVA on accuracy was re-run, certain differences were apparent in comparison to the results from the full sample. Most importantly, the interaction between source expertise and syllogism type had emerged as marginally significant. Hence, these potentially problematic participants were somewhat less likely to show the source expertise bias compared to the full sample. However, it is difficult to determine with certainty whether all of these participants simply did not pay attention to the instructions of the reasoning task or whether some of them were legitimately less influenced by the source expertise information. Nevertheless, the pattern of results supports the likelihood that some of these participants did not complete the overall experiment properly.

The lack of source expertise effect in the full sample of the present experiment may nonetheless be accounted for through the literature on source credibility and individual differences. Much previous research has had difficulty finding source credibility effects because argument quality and other content-related aspects of a message have often overshadowed the source of the information (Petty & Wegener, 1998). For example, Sparks and Rapp (2011) conducted four experiments and found little evidence of a source credibility effect. In each subsequent experiment, these authors had presented participants with increasing amounts of

written information about the sources and encouragement to consider this information while making decisions. It was only in the fourth experiment in which participants were explicitly told to evaluate the sources that their reading latencies showed an influence due to the source information. Moreover, the sources in Sparks and Rapp's experiment were relevant to people's processing of the text, whereas in the present experiment the source information was completely irrelevant. Hence, although the present results are not completely consistent with Copeland et al. (2011), the lack of source expertise bias is understandable in consideration of the broader literature. Furthermore, the relationship of this bias to people's CRT scores indicates that the degree of source expertise effect varies in accordance with people's thinking characteristics. Thus, it is not surprising that the bias is small because it tends to only occur among a sub-sample of the general population.

Prior research had found that the size of source credibility effects varies in accordance with different media of presentation (Wilson & Sherrell, 1993). For instance, two studies showed that likeable or trustworthy sources had a greater influence through video, whereas unlikeable or untrustworthy sources had a greater influence through written information (Chaiken & Eagly, 1983; Andreoli & Worchel, 1978). Although these studies involved different traits than expertise and used videos rather than pictures, it is plausible that similar results occurred in the present study. Unfortunately, however, it is not as simple to identify such a pattern within a syllogistic reasoning paradigm. In a persuasion paradigm, the outcome is often a simple judgement of one's degree of agreement with a source. In contrast, the outcome of showing the source expertise bias in one's accuracy performance does not allow for a distinction between the influences of the expert versus the non-expert. That is, showing the bias might represent either people's tendency to accept the expert or to reject the non-expert (or both). Moreover, if people accept experts more

in the visual medium yet accept non-experts more in the written medium, then the effect of an interaction between medium and source might have been masked due to the fact that the bias occurs oppositely in the valid versus invalid syllogisms. Thus, further research might benefit from including a baseline condition containing no source information in order that the comparison between accuracy on the expert-provided syllogisms versus accuracy on syllogisms provided by the non-expert would be more interpretable across media of presentation.

Furthermore, Chaiken and Eagly (1976) found that the effect of the medium of presentation depended on the difficulty level of the message content. That is, when the message was easy to understand, they found that persuasion was highest through the video medium although comprehension was similar across mediums. However, when the message was difficult to understand, then persuasion and comprehension were highest through the written medium. Participants have difficulty evaluating syllogisms. Thus there is a possibility that the addition of source information (particularly that of the visual medium) in the present experiment made it even more difficult for people to evaluate the syllogisms, resulting in lower accuracy in people who viewed the visual information. However, this effect would be difficult to distinguish from the fact that the numbers of high CRT scorers were not equal across the two media. Regardless, further experimentation that includes message content with distinct levels of difficulty might be helpful to differentiate the effects of using visual images versus using strictly written information. For example, an experiment could test how people respond to syllogisms presented using different media by creating groups of syllogisms according to the number of mental models required. Alternatively, the difficulty level of understanding the content within syllogisms could be altered, while still maintaining neutrality.

More importantly, additional consideration of the construct represented by the CRT is clearly essential because of the consistent results indicating that this measure is strongly related to higher order cognitive abilities (such as reasoning and intelligence). The CRT is a relatively new measure of cognitive ability; hence it is likely that further research will suggest some adjustment and increased precision in defining the cognitive construct that it represents. According to Frederick (2005), the CRT measures "...the ability or disposition to resist reporting the response that first comes to mind" (p. 35). However, Frederick's definition does not include mention of the problem-solving processes that need to occur after the individual has resisted the initial response. Although for some people this further problem-solving process is likely a natural extension of their thinking after they have rejected the incorrect response, for others there is likely a tendency to experience some sort of further obstacle to solving the problem. That is, due to the prominence of the intuitive response that comes to mind initially, people will be likely to experience some of the common barriers to problem-solving. For example, people might be inhibited by the experience of a "mental set" or "unnecessary constraints", tendencies which cause people to employ incorrect methods in problem-solving due to their past experience that a technique had worked for similar problems or due to their mistaken placement of boundaries that cause fixation on an incorrect technique, respectively (Weiten & McCann, 2010). Thus, the persistence of such flawed methods would then make people unable to employ the proper problem-solving skills needed to determine the correct answer.

In summary, several characteristics may be related to success on the CRT problems. These characteristics need to be outlined and tested thoroughly in future experiments. In particular, there are at least two categories of abilities involved in solving the problems. Firstly, because the three questions on the CRT could be classified as basic mathematical word

problems, the ability to perform mathematical operations should be critical in determining the answers. Accordingly, Campitelli and Gerrans (2013) found that mathematical skill as measured by a numeracy test explained significant variability in CRT scores. Hence, instruction and practice in mathematics will clearly be helpful toward the accurate completion of the CRT.

Nonetheless, it is apparent that there is more to solving the CRT problems than mathematical ability. The three questions are uniquely related to each other because of the way that they all quickly evoke intuitive wrong answers which must be ignored before further calculation can proceed to find the correct answers. Hence, another obvious factor involved in solving the problems is the particular thinking style that the test purports to measure. That is, people's tendency to inhibit their intuitive responses and to think about alternative solutions is expected to be crucial to their ability to score highly on the CRT. This tendency is considered to involve reflective thinking as well as other processes of System 2 or executive functioning. For instance, measures of rational thinking ability (to a high degree) and actively open-minded thinking (to a lesser degree) have been found to explain significant variability in CRT scores (Campitelli & Gerrans, 2013). Still, further research will be needed to more thoroughly refine understanding of the characteristics that the CRT represents. Campitelli and Gerrans tested each of the three problems individually and found that the "bat-and-ball" problem was most related to inhibition and least related to mathematical ability. Hence, they suggested that similar, additional problems should be found based on whether they also emphasize the inhibition factor rather than the mathematics factor. Accordingly, it has recently been suggested that the number of items on the CRT test should be expanded for reasons of increasing its consistency and to counter the fact that the test is becoming less useful as more people become familiar with it (Toplak, West, & Stanovich, 2014).

A third factor that may be involved in solving the CRT problems is accuracy motivation, which is a basic element of both the ELM and the HSM theories of attitude change. Many of the attitude change studies that support these two models used situational manipulations to increase or decrease people's degree of accuracy motivation. Such manipulations influenced people's perceptions of the task importance or the personal relevance of the issue at hand, which typically caused them to exert more or less mental effort accordingly. When people have high accuracy motivation, they are more likely to take additional time to come up with the most appropriate answer. Likewise, participants who scored high on the CRT in the present experiment were more likely to take additional time to solve the syllogisms. Specifically, both people who score highly on the CRT and people who have greater accuracy motivation are more likely to exert high effort in the processing of information and less likely to use heuristics such as source expertise. Hence, it is logical to consider that having high accuracy motivation is closely related to having a high CRT score. Therefore, in addition to measuring mathematical skill and reflective thinking style, it is possible that the CRT is also capturing people's dispositional level of accuracy motivation to some degree. Moreover, the present finding involving the relationship between CRT scores and source expertise bias corresponds well with the frameworks of the two most classic theories in persuasion literature.

Furthermore, it is noteworthy that people who value their accuracy in a context such as the present experiment, in which they receive no consequence or feedback for their responses, are likely to be people who have a greater intrinsic desire to perform well on all cognitive tasks. Therefore, it is not surprising that these people include only a relatively small sub-sample of the overall university student population, especially considering that people in general have strong tendencies to be cognitive misers and thus to prioritize their exertion of mental effort

accordingly. In fact, the CRT has been referred to as a measure on which low scores indicate a greater degree of “miserly processing” (Toplak et al., 2011).

Accordingly, future research would likely benefit from distinguishing more clearly between the elements of motivation versus ability, which both seem to be measured by the CRT. As suggested strongly by the HSM, these two elements will influence people’s judgements in many situations. Specifically, if both motivation and ability are high, then people will be more likely to engage in systematic processing. On the other hand, if either motivation or ability is low, then people will be more likely to rely on heuristic processing. Hence, it seems likely that both motivation and ability are particularly critical in regard to understanding the CRT. That is, people who score high on the CRT likely have both high motivation and ability, which also makes them less likely to use irrelevant heuristic cues, more likely to take additional time to consider alternative solutions, and thus overall more likely to perform accurately. Interestingly, it seems that the elements of motivation and ability might also coincide well with Stanovich’s (2012) distinction between the reflective mind and the algorithmic mind, respectively. Thus, although motivation might provide a critical first step in determining the nature of a problem, ability is imperative to continue the thinking process effectively. Moreover, recall that the CRT and the AOT scale were included in the present study in order to represent operations of the reflective mind, whereas WM was included to represent processes of the algorithmic mind. However, the CRT measures both reflective and algorithmic operations, although the former are considered to be the central purpose for the measurement. Furthermore, the inclusion of both reflective and algorithmic aspects in a hierarchical manner as outlined in Stanovich’s model might turn out to be the chief explanatory factor regarding why the CRT has shown a serious advantage over other popular cognitive measures (e.g., intelligence tests, working memory, and

the NFC scale; Frederick, 2005; Toplak et al., 2011) in predicting people's performance in many higher order decision-making tasks.

Further consideration of people's thinking dispositions in relation to higher order cognitive abilities should also be carried out in regard to the AOT measure. Specifically, determining whether the inconsistent findings with this measure are empirically important should be a goal of future studies. Although overall test reliability of the AOT measure was good (i.e., Cronbach's  $\alpha = .84$ ), there is still the possibility that the AOT measure contains certain flaws. For instance, Marsh and Pastor (2011) conducted a factor analysis on the 41 items of the AOT questionnaire using data from 445 participants. Rather than finding a 6-factor solution (corresponding to the number of subscales) or a unidimensional solution (corresponding to the standard method of scoring the items), a 4-factor solution emerged. More importantly, the inter-item correlations were low, a relatively large number of items were problematic (i.e., they failed to load on the factors), the item communalities were low, and the total percentage of variance accounted for was low. Overall, this indicates that the items might be measuring distinct concepts and that there is likely confusing or ambiguous wording which is not interpreted equivalently by different participants. Hence, the AOT measure might benefit from revision or removal of problematic items. Marsh and Pastor concluded that the AOT test does not represent the proposed construct appropriately. Some of the included components might not be essential elements of the construct, but rather are only related to it. Although this single factor analysis is not adequate to discount previous findings with the AOT scale, future research would clearly be helpful to further investigate the AOT measure and perhaps contribute to refining the scale so that it better represents the desired construct.

If the AOT measure is shown to be flawed, the most important question from the perspective of the present experiment becomes: Are the cognitive constructs that the AOT scale measures able to explain the present results? Marsh and Pastor's (2011) results indicate that some people interpret items on the AOT scale in an inconsistent manner. Hence, for instance, perhaps a portion of the high scoring participants are people with a tendency to accept new information exactly as it is and to rapidly reform their existing beliefs accordingly, rather than more generally questioning all information (both new and existing) in a critical manner. Thus in this case people who score highly might actually be those who think *less* carefully about information. This type of result would represent the opposite of the construct that the AOT scale is intended to measure. Therefore the occurrence of this pattern would explain why results might be inconsistent across studies. In contrast, as previously mentioned, the present result which indicates that higher AOT scorers show higher source expertise bias might be partially due to socially desirable responding. That is, people who responded to the AOT items by indicating that they were particularly likely to think carefully about information either were purposely attempting to appear more thoughtful or had the mistaken belief that they actually were more thoughtful. Therefore, future research might benefit from including a measure of social desirability as a covariate in order to control for people's tendency to respond in a socially desirable manner. For this purpose, there are two popular measures which might be useful: the Marlowe-Crowne Social Desirability Scale (Crowne & Marlow, 1960) or the Balanced Inventory of Desirable Responding (Paulhus, 1991).

The results of this thesis have implications for an array of future research questions. Firstly, the findings suggest that new samples of participants might be equally as likely to show the source credibility effect that was found in Copeland et al. (2011) as they would be to not

show the effect as in the present experiment. Thus, additional replications would be useful to better understand the phenomenon. Secondly, including individual differences measures in future tests of the source credibility effect will be particularly informative. The CRT seems to be the most exciting candidate measurement for helping to understand why people differ in their higher order cognitive abilities such as reasoning and decision-making. Consequently, the CRT must be explored with increasing empirical and theoretical depth. Nonetheless, AOT and WM will also require more consideration within the source credibility and syllogistic reasoning paradigm. In sum, people's individual differences in thinking dispositions appear to have a central influence on their ability to avoid being biased by source expertise information while they engage in formal reasoning.

## References

- Andreoli, V., & Worchel, S. (1978). Effects of Media, Communicator, and Message Position on Attitude Change. *Public Opinion Quarterly*, 42(1), 59-70.
- Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Science*, 4(11), 417-423.
- Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 8, pp. 47-89). Academic Press.
- Ball, L. J., Phillips, P., Wade, C. N., & Quayle, J. D. (2006). Effects of belief and logic on syllogistic reasoning: Eye-movement evidence for selective processing models. *Experimental Psychology*, 53, 77-86.
- Begg, I., & Denny, J. P. (1969). Empirical reconciliation of atmosphere and conversion interpretations of syllogistic reasoning. *Journal of Experimental Psychology*, 81, 351-354.
- Braine, M. D. S., & Rumin, B. (1983). Logical reasoning. In J. H. Flavell & E. M. Markman (Eds.), *Carmichael's handbook of child psychology, Vol. 3. Cognitive development* (4th ed., pp. 263-339). New York, NY: Wiley.
- Bromme, R., Kienhues, D., & Porsch, T. (2011). Who knows what and who can we believe? Epistemological beliefs are beliefs about knowledge (mostly) to be attained from others. In L. D. Bendixen & F. C. Haerle (Eds.), *Personal Epistemology in the Classroom: Theory, Research, and Implications for Practice*. Cambridge: Cambridge University Press.
- 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, 805-818.
- Campitelli, G., & Gerrans, P. (2013). Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach. *Memory and Cognition*, N/A(N/A), p14, USA, DOI: 10.3758/s13421-013-0367-9.
- Case, R., Kurland, M. D., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, 33, 386-404.
- Chaiken, S. (1978). The use of source versus message cues in persuasion: An information processing analysis (Doctoral dissertation, University of Massachusetts-Amherst). *Dissertation Abstracts International*, 39, 438B.
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39, 752-766.
- Chaiken, S., & Eagly, A. H. (1976). Communication modality as a determinant of message persuasiveness and message comprehensibility. *Journal of Personality and Social Psychology*, 34(4), 605-614.
- Chaiken, S., & Eagly, A. H. (1983). Communication modality as a determinant of persuasion: The role of communicator salience. *Journal of Personality and Social Psychology*, 45, 241-256.
- Chaiken, S., & Ledgerwood, A. (2012). A theory on heuristic and systematic information processing. In P. A. M. Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of Theories of Social Psychology*, Vol. 1 (pp. 224-245). London: Sage.

- Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic processing within and beyond the persuasion context. In J.S. Uleman & J.A. Bargh (Eds.) *Unintended Thought* (pp. 212-252). New York: Guilford Press.
- Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgement. *Journal of Personality and Social Psychology*, 66(3), 460-473.
- Chapman, L. J., & Chapman, J. P. (1959). Atmosphere effect re-examined. *Journal of Experimental Psychology*, 58, 220–226.
- Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, 38, 191–258.
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769-786.
- Copeland, D. E., Gunawan, A., & Bies-Hernandez, N. J. (2011). Source credibility and syllogistic reasoning. *Memory and Cognition*, 39, 117-127.
- Copeland, D. E., & Radvansky, G. A. (2004). Working memory and syllogistic reasoning. *The Quarterly Journal of Experimental Psychology*, 57A(8), 1437–1457.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory*. Odessa, FL: Psychological Assessment Resources.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4), 349-354.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, 19, 450-466.

- Demeure, V., Bonnefon, J.-F., & Raufaste, E. (2009). Politeness and conditional reasoning: Interpersonal cues to the indirect suppression of deductive inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(1), 260-266.
- Engle, R. (2005). Reading Span (E-Prime). The Attention and Working Memory Lab, Georgia Institute of Technology. Retrieved December 12, 2006, from <http://psychology.gatech.edu/renglelab/>.
- Epstein, S., & Meier, P. (1989). Constructive thinking: A broad coping variable with specific components. *Journal of Personality and Social Psychology*, 57, 332-350.
- Erickson, J. R. (1974). A set analysis theory of behavior in formal syllogistic reasoning tasks. In R. Solso (Ed.), *Loyola symposium on cognition* (Vol. 2, pp. 305–330). Hillsdale, NJ: Erlbaum.
- Erwin, T. D. (1981). *Manual for the Scale of Intellectual Development*. Harrisonburg, VA: Developmental Analytics.
- Erwin, T. D. (1983). The Scale of Intellectual Development: Measuring Perry's scheme. *Journal of College Student Personnel*, 24, 6-12.
- Evans, J. St. B. T. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin*, 128, 978–996.
- Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences*, 7, 454–459.
- Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition*, 11, 295–306.
- Evans, J. St. B. T., Handley, S. J., Harper, C. N. J., & Johnson-Laird, P. N. (1999). Reasoning

- about necessity and possibility: A test of the mental model theory of deduction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1495–1513.
- Ford, M. (1995). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition*, 54, 1–71.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19, 25–42.
- French, L., Garry, M., & Mori, K. (2011). Relative – not absolute – judgments of credibility affect susceptibility to misinformation conveyed during discussion. *Acta Psychologica*, 136, 119–128.
- Geurts, B. (2003). Reasoning with quantifiers. *Cognition*, 86, 223–251.
- Gigerenzer, G., & Brighton, H. (2009). Homo Heuristicus: Why Biased Minds Make Better Inferences. *Topics in Cognitive Science*, 1(1), 107-143.
- Gilovich, T. & Griffin, D. (2002). Introduction – Heuristics and Biases: Then and Now. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (pp. 1-18). Cambridge, UK: Cambridge University Press.
- Goldacre, B. (2010). *Bad Science*. Toronto, Ontario: McClelland & Stewart Ltd.
- Gordon, R., & Bryant, K. (2010). *You don't act like you trust me: Dissociations between behavioral and explicit measures of source credibility judgment*. Manuscript submitted for publication.
- Homer, P. M., & Kahle, L. R. (1990). Source Expertise, Time of Source Identification, and Involvement in Persuasion: An Elaborative Processing Perspective. *Journal of Advertising*, 19(1), 30-39.
- Hovland, C. I., & Weiss, W. (1951). The Influence of Source Credibility on Communication

- Effectiveness. *Public Opinion Quarterly*, 15(4), 635-650.
- Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition*, 16, 1-61.
- Johnson-Laird, P. N. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness*. Cambridge: Harvard University.
- Johnson-Laird, P. N. (1975). Models of deduction. In R. Falmagne (Ed.), *Reasoning: Representation and process* (pp. 7-54). Springdale, NJ: Erlbaum.
- Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.
- Kahneman, D., & Frederick, S. (2005). A Model of Heuristic Judgement. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 67-293). New York: Cambridge University Press.
- Kane, M. J., & Engle, R. W. (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.
- Kiesler, S. B., & Mathog, R. B. (1968). Distraction Hypothesis in Attitude Change: Effects of Effectiveness. *Psychological Reports*, 23, 1123-1133.
- Kilpatrick, S. G., Manktelow, K. I., & Over, D. E. (2007). Power of source as a factor in deontic reasoning. *Thinking and Reasoning*, 13(3), 295-317.
- Khemlani, S., & Johnson-Laird, P. N. (2012). Theories of the Syllogism: A Meta-Analysis. *Psychological Bulletin*, 138(3), 427-457.
- Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). London: Routledge.
- Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, 28(4), 612-625.
- Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2011). Individual

- differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making*, 25, 361–381.
- Lilienfeld, S. O., Lynn, S. J., Ruscio, J., & Beyerstein, B. L. (2010). *50 Great Myths of Popular Psychology: Shattering Widespread Misconceptions*. United Kingdom: Wiley-Blackwell.
- Macpherson, R., & Stanovich, K. E. (2007). Cognitive ability, thinking dispositions, and instructional set as predictors of critical thinking. *Learning and Individual Differences*, 17, 115-127.
- Maddux, J. E., & Rogers, R. W. (1980). Effects of source expertness, physical attractiveness, and supporting arguments on persuasion: a case of brains over beauty. *Journal of Personality and Social Psychology*, 39, 235-244.
- Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition*, 17(1), 11-17.
- Marsh, K. R., & Pastor, D. A. (2011). Stanovich and West's (2007) Actively Open-Minded Thinking Scale: An Examination of Factor Structure. *APS annual meeting*, Washington, D.C.
- McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The Relationship Between Working Memory Capacity and Executive Functioning: Evidence for a Common Executive Attention Construct. *Neuropsychology*, 24(2), 222-243.
- Mechanic (Isolated on White) – Stock Image [Online image]. Retrieved February 4, 2013 from <http://www.istockphoto.com/stock-photo-21886396-mechanic-isolated-on-white.php?st=6cd96d0>
- Newell, A. (1981). Reasoning, problem solving and decision processes: The problem space as a fundamental category. In R. Nickerson (Ed.), *Attention and Performance VIII* (pp. 693–718). Hillsdale, NJ: Erlbaum.

- Newstead, S.E., Pollard, P., Evans, J. St.B.T., & Allen, J.L. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition*, 45, 257-284.
- Nguyen, H., & Masthoff, J. (2007). Is it me or is it what I say? Source image and persuasion. In Y. de Kort, W. IJsselsteijn, C. Midden, B. Eggen, & B. J. Fogg (Eds.), *Lecture Notes in Computer Science: Vol. 4744. Persuasive Technology* (pp. 231-242). Berlin, Germany: Springer-Verlag.
- Norman, D. A., & Shallice, T. (1980). Attention to Action: Willed and Automatic Control of Behavior. *Center for Human Information Processing (CHIP) Technical Report 99*. San Diego: University of California.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.
- Peirce, J. W. (2007). PsychoPy – Psychophysics Software in Python. *Journal of Neuroscience Methods*, 162(1-2), 8-13.
- Peirce, J. W. (2008). Generating Stimuli for Neuroscience Using PsychoPy. *Frontiers in Neuroinformatics*, 2(10).
- Petty, R. E., & Briñol, P. (2012). The elaboration likelihood model. In P. A. M. Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of Theories of Social Psychology, Vol. 1* (pp. 224-245). London: Sage.
- Petty, R. E., & Cacioppo, J. T. (1979). Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology*, 37, 1915-1926.
- Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and Persuasion: Classic and Contemporary*

- Approaches*. Dubuque, IA: Wm. C. Brown. Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Advances in Experimental Social Psychology, Volume 19* (pp. 123-205). Orlando, Florida: Academic Press, Inc.
- Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, 41(5), 847-855.
- Petty, R. E., Haugtvedt, C., & Smith, S. M., (1995). Elaboration as a determinant of attitude strength: Creating attitudes that are persistent, resistant, and predictive of behavior. In R.E. Petty and J.A. Krosnick (eds), *Attitude Strength: Antecedents and Consequences*, pp. 93-130. Mahwah, NJ: Erlbaum.
- Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed.). New York: McGraw-Hill.
- Petty, R. E., Wells, G. L., & Brock, T. C. (1976). Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and Social Psychology*, 34, 874-884.
- Polk, T. A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review*, 102, 533–566.
- Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' research. *Journal of Applied Social Psychology*, 34(2), 243-281.
- Porter, S., Brinke, L., & Gustaw, C. (2010). Dangerous decisions: the impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime and Law*, 16(6), 477-491.

Professor Series (XXL) – Stock Image [Online image]. Retrieved February 4, 2013 from

<http://www.istockphoto.com/stock-photo-8579621-professor-series-xxl.php?st=7c84570>

Quayle, J. D., & Ball, L. J. (2000). Working memory, metacognitive uncertainty, and belief bias in syllogistic reasoning. *The Quarterly Journal of Experimental Psychology*, 53A(4), 1202-1223.

Raven, B. H., & French, J. R. P. (1958). Legitimate Power, Coercive Power, and Observability in Social Influence. *Sociometry*, 21(2), 83-97.

Reinhard, M.-A., Messner, M., & Sporer, S. L. (2006). Explicit Persuasive Intent and Its Impact on Success at Persuasion – The Determining Roles of Attractiveness and Likeableness. *Journal of Consumer Psychology*, 16(3), 249-259.

Revlis, R. (1975). Two models of syllogistic reasoning: Feature selection and conversion. *Journal of Verbal Learning and Verbal Behavior*, 14, 180–195.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA: MIT Press.

Rokeach, M. (1960). *The open and closed mind*. New York: Basic Books.

Sà, 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.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002a). E-Prime User's Guide. Pittsburgh: Psychology Software Tools Inc.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002b). E-Prime Reference Guide. Pittsburgh: Psychology Software Tools Inc.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological*

- Review*, 63(2), 129-138.
- Sparks, J. R., & Rapp, D. N. (2011). Readers' reliance on source credibility in the service of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 230–247.
- Stanovich, K. E. (2004). *The Robot's Rebellion: Finding Meaning in the Age of Darwin*. Chicago, IL: The University of Chicago Press.
- Stanovich, K. E. (2009). *What Intelligence Tests Miss: The Psychology of Rational Thought*. New Haven, CT: Yale University Press.
- Stanovich, K. E. (2012). On the Distinction Between Rationality and Intelligence: Implications for Understanding Individual Differences in Reasoning. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning*. New York: Oxford University Press, Inc. Chapter 22, 342-365.
- 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, 342–357.
- Stenning, K., & Yule, P. (1997). Image and language in human reasoning: A syllogistic illustration. *Cognitive Psychology*, 34, 109–159.
- Stevenson, R. J., & Over, D. E. (2001). Reasoning from uncertain premises: Effects of expertise and conversational context. *Thinking and Reasoning*, 7(4), 367-390.
- Thanh Ha, T. (2012, April 27). Nutella backs off 'balanced breakfast' claims after hit with \$3-million lawsuit. *Globe and Mail*. Retrieved from <http://www.theglobeandmail.com>
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a

- predictor of performance on heuristics-and-biases tasks. *Memory and Cognition*, 39, 1275-1289.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking and Reasoning*, 20(2), 147-168.
- Troidahl, V., & Powell, F. (1965). A short-form dogmatism scale for use in field studies. *Social Forces*, 44, 211-215.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language*, 28, 127-154.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124-1131.
- Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience*, 20(3), 470-477.
- Wetherick, N. E., & Gilhooly, K. J. (1995). "Atmosphere," matching, and logic in syllogistic reasoning. *Current Psychology*, 14, 169-178.
- Wilson, E. J., & Sherrell, D. L. (1993). Source Effects In Communication and Persuasion Research: A Meta-Analysis of Effect Size. *Journal of the Academy of Marketing Science*, 21(2), 101-112.
- Woodworth, R. S., & Sells, S. B. (1935). An Atmosphere Effect in Formal Syllogistic Reasoning. *Journal of Experimental Psychology*, 18(4), 451-460.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9, Monograph Suppl. No. 2, part 2, 1-27.

## Appendix A

Structure of the 32 syllogisms: mood/figure of two premises (e.g., AA1) and mood/order of corresponding conclusion (e.g., Aac). The necessary constitute correct (i.e., valid) syllogisms, while the possible strong constitute incorrect (i.e., invalid) syllogisms.

| Necessary syllogisms |     | Possible Strong syllogisms |     |
|----------------------|-----|----------------------------|-----|
| AA1                  | Aac | AO1                        | Oac |
| AE1                  | Eac | II1                        | Iac |
| AI2                  | Iac | AO2                        | Oac |
| IE2                  | Oac | IA2                        | Iac |
| AE3                  | Eac | IA3                        | Iac |
| EA3                  | Eac | IO3                        | Oac |
| AI4                  | Iac | AA4                        | Aac |
| AE4                  | Oac | OA4                        | Oac |
| IA1                  | Ica | AI1                        | Ica |
| EA1                  | Oca | IO1                        | Oca |
| AA2                  | Aca | II2                        | Ica |
| AI2                  | Ica | IO2                        | Oca |
| EA3                  | Eca | AI3                        | Ica |
| EI3                  | Oca | IA3                        | Ica |
| A14                  | Ica | IO4                        | Oca |
| EI4                  | Oca | OI4                        | Oca |

## Appendix B

Full set of syllogisms with neutral content (hobbies and status terms). The items that were altered from Copeland et al.'s (2011) original stimuli are shown here in italics for ease of identification.

| <b>Premise 1</b>                                  | <b>Premise 2</b>                                | <b>Conclusion</b>                              |
|---|---|--|
| All Mac users are painters                        | All painters are bilinguals                     | All Mac users are bilinguals                   |
| All dog owners are recyclers                      | All cigar smokers are dog owners                | All cigar smokers are recyclers                |
| All bilinguals are cyclists                       | No racquetball players are cyclists             | No bilinguals are racquetball players          |
| All Mac users are bilinguals                      | Some tennis players are Mac users               | Some bilinguals are tennis players             |
| All bakers are wine drinkers                      | Some bakers are swimmers                        | Some wine drinkers are swimmers                |
| All cyclists are boaters                          | Some cyclists are wine drinkers                 | Some wine drinkers are boaters                 |
| No <i>university</i> students are musicians       | All readers are musicians                       | No readers are <i>university</i> students      |
| Some married people are skiers                    | All skiers are tennis players                   | Some tennis players are married people         |
| All dog owners are chess players                  | All dog owners are <i>Harper</i> supporters     | All chess players are <i>Harper</i> supporters |
| All coffee drinkers are boaters                   | Some boaters are married people                 | Some married people are coffee drinkers        |
| All tennis players are <i>university</i> students | Some swimmers are <i>university</i> students    | Some swimmers are tennis players               |
| All racquetball players are recyclers             | Some recyclers are not cyclists                 | Some racquetball players are not cyclists      |
| All bakers are campers                            | Some wine drinkers are not bakers               | Some campers are not wine drinkers             |
| Some coffee drinkers are married people           | All dog owners are married people               | Some coffee drinkers are dog owners            |
| Some coffee drinkers are musicians                | Some readers are coffee drinkers                | Some readers are musicians                     |
| Some Mac users are skiers                         | Some Mac users are not <i>Harper</i> supporters | Some <i>Harper</i> supporters are not skiers   |
| All musicians are bilinguals                      | No bilinguals are cigar smokers                 | No musicians are cigar smokers                 |
| All dog owners are readers                        | No dog owners are skiers                        | Some readers are not skiers                    |
| All bakers are skiers                             | Some racquetball players are bakers             | Some racquetball players are skiers            |
| No painters are <i>Harper</i> supporters          | All <i>Harper</i> supporters are chess players  | Some chess players are not painters            |
| No boaters are married people                     | All <i>Harper</i> supporters are married people | No boaters are <i>Harper</i> supporters        |

|   |   |  |
|---|---|--|
| No cigar smokers are musicians                              | Some football fans are musicians                  | Some football fans are not cigar smokers   |
| No swimmers are football fans                               | Some swimmers are wine drinkers                   | Some wine drinkers are not football fans   |
| Some cyclists are hospital volunteers                       | No campers are cyclists                           | Some hospital volunteers are not campers   |
| Some boaters are hospital volunteers                        | All chess players are boaters                     | Some hospital volunteers are chess players |
| Some bakers are recyclers                                   | All hospital volunteers are recyclers             | Some hospital volunteers are bakers        |
| Some campers are football fans                              | Some football fans are cigar smokers              | Some campers are cigar smokers             |
| Some hospital volunteers are <i>university</i> students     | Some <i>university</i> students are not recyclers | Some recyclers are not hospital volunteers |
| Some tennis players are chess players                       | Some Mac users are not tennis players             | Some Mac users are not chess players       |
| Some painters are football fans                             | Some readers are not football fans                | Some painters are not readers              |
| Some <i>university</i> students are not racquetball players | All <i>university</i> students are swimmers       | Some racquetball players are not swimmers  |
| Some coffee drinkers are not painters                       | Some coffee drinkers are campers                  | Some campers are not painters              |

## Appendix C

Actively Open-minded Thinking (AOT) measurement comprised of six subscales.

|    |   |
|----|---|
| 1  | Even though freedom of speech for all groups is a worthwhile goal, it is unfortunately necessary to restrict the freedom of certain political groups. (R) |
| 2  | What beliefs you hold have more to do with your own personal character than the experiences that may have given rise to them. (R)                         |
| 3  | I tend to classify people as either for me or against me. (R)   |
| 4  | A person should always consider new possibilities.  |
| 5  | There are two kinds of people in this world: those who are for the truth and those who are against the truth. (R)   |
| 6  | Changing your mind is a sign of weakness. (R)   |
| 7  | I believe we should look to our religious authorities for decisions on moral issues. (R)  |
| 8  | I think there are many wrong ways, but only one right way, to almost anything. (R)  |
| 9  | It makes me happy and proud when someone famous holds the same beliefs that I do. (R)   |
| 10 | Difficulties can usually be overcome by thinking about the problem, rather than through waiting for good fortune.   |
| 11 | There are a number of people I have come to hate because of the things they stand for. (R)  |
| 12 | Abandoning a previous belief is a sign of strong character.   |
| 13 | No one can talk me out of something I know is right. (R)  |
| 14 | Basically, I know everything I need to know about the important things in life. (R)   |
| 15 | It is important to persevere in your beliefs even when evidence is brought to bear against them. (R)  |
| 16 | Considering too many different opinions often leads to bad decisions. (R)   |
| 17 | There are basically two kinds of people in this world, good and bad. (R)  |
| 18 | I consider myself broad-minded and tolerant of other people's lifestyles.   |
| 19 | Certain beliefs are just too important to abandon no matter how good a case can be made against them. (R)   |
| 20 | Most people just don't know what's good for them. (R)   |
| 21 | It is a noble thing when someone holds the same beliefs as their parents. (R)   |
| 22 | Coming to decisions quickly is a sign of wisdom. (R)  |
| 23 | I believe that loyalty to one's ideals and principles is more important than "open-mindedness." (R)   |
| 24 | Of all the different philosophies which exist in the world there is probably only one which is correct. (R)   |
| 25 | My beliefs would not have been very different if I had been raised by a different set of parents. (R)   |
| 26 | If I think longer about a problem I will be more likely to solve it.  |
| 27 | I believe that the different ideas of right and wrong that people in other societies have may be valid for them.  |
| 28 | Even if my environment (family, neighborhood, schools) had been different, I probably would have the same religious views. (R)                            |
| 29 | There is nothing wrong with being undecided about many issues.  |
| 30 | I believe that laws and social policies should change to reflect the needs of a changing  |

|    |   |
|----|---|
|    | world.  |
| 31 | My blood boils over whenever a person stubbornly refuses to admit he's wrong. (R)                                     |
| 32 | I believe that the "new morality" of permissiveness is no morality at all. (R)  |
| 33 | One should disregard evidence that conflicts with your established beliefs. (R)                                       |
| 34 | Someone who attacks my beliefs is not insulting me personally.  |
| 35 | A group which tolerates too much difference of opinion among its members cannot exist for long. (R)                   |
| 36 | Often, when people criticize me, they don't have their facts straight. (R)  |
| 37 | Beliefs should always be revised in response to new information or evidence.  |
| 38 | I think that if people don't know what they believe in by the time they're 25, there's something wrong with them. (R) |
| 39 | I believe letting students hear controversial speakers can only confuse and mislead them. (R)                         |
| 40 | Intuition is the best guide in making decisions. (R)  |
| 41 | People should always take into consideration evidence that goes against their beliefs.                                |