2013

Motivated Reasoning and Response Bias: A Signal Detection Approach

Trippas, Dries

http://hdl.handle.net/10026.1/2853

http://dx.doi.org/10.24382/4784
University of Plymouth

School of Psychology
Faculty of Science and Technology
Plymouth University

Motivated Reasoning and Response Bias: A Signal Detection Approach

By Dries Trippas

A thesis submitted to Plymouth University in partial fulfilment for the degree of

Doctor of Philosophy

July 2013

Abstract

Motivated Reasoning and Response Bias: A Signal Detection Approach

Dries Trippas

The aim of this dissertation was to address a theoretical debate on belief bias. Belief bias is the tendency for people to be influenced by their prior beliefs when engaged in deductive reasoning. Deduction is the act of drawing necessary conclusions from premises which are meant to be assumed as true. Given that the logical validity of an argument is independent of its content, being influenced by your prior beliefs in such content is considered a bias. Traditional theories posit there are two belief bias components. Motivated reasoning is the tendency to reason better for arguments with unbelievable conclusions relative to arguments with believable conclusions. Response bias is the tendency to accept believable arguments and to reject unbelievable arguments. Dube et al. (2010) pointed out critical methodological problems that undermine evidence for traditional theories. Using signal detection theory (SDT), they found evidence for response bias only. We adopted the SDT method to compare the viability of the traditional and the response bias accounts. In Chapter 1 the relevant literature is reviewed. In Chapter 2 four experiments which employed a novel SDT-based forced choice reasoning method are presented, showing evidence compatible with motivated reasoning. In Chapter 3 four experiments which used the receiver operating characteristic (ROC) method are presented. Crucially, cognitive ability turned out to be linked to motivated reasoning. In Chapter 4 three experiments are presented in which we investigated the impact of cognitive ability and analytic cognitive style on belief bias, concluding that cognitive style mediated the effects of cognitive ability on motivated reasoning. In Chapter 5 we discuss our findings in light of a novel individual differences account of belief bias. We conclude that using the appropriate measurement method and taking individual differences into account are two key elements to furthering our understanding of belief bias, human reasoning, and cognitive psychology in general.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisors Prof. Simon Handley and Dr. Michael Verde for their great guidance and for teaching me how to do research.

I would also like to thank my parents Ivan and Anne, and my brothers Wout and Bavo for all their support.

Thanks to the following people for entertaining me throughout the process of doing this PhD: Koen Vanbrabant, An Thys, David West, Olli Lattimer, Tim Rumbell, James Humble and Laura Koenig

Finally, huge thanks go out to Gord Pennycook for running the final experiment. It really tied the dissertation together.

Author's Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award. Work submitted for this research degree at Plymouth University has not formed part of any other degree either at Plymouth University or at another establishment.

This study was financed with the aid of a studentship from the School of Psychology. The thesis is the result of the author's own investigations, carried out under the guidance of his supervisory team. Other sources are explicitly referenced.

Publications:

Trippas, D., Handley, S. J., & Verde, M. F. (2013). The SDT Model of Belief Bias: Complexity, Time and Cognitive Ability Mediate the Effects of Believability. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* doi:10.1037/a0032398

Trippas, D., Pennycook, G., Verde, M. F., & Handley, S. J. (under review). Better but More Biased: Analytic Cognitive Style Predicts Motivated Reasoning. *Psychonomic Bulletin & Review*

Trippas, D., Verde, M. F., & Handley, S. J. (in preparation). Assessing Belief Bias Using Forced Choice Reasoning: Evidence for Motivated Reasoning and Individual Differences.

Conference Presentations:

Trippas, D., Handley, S. J., & Verde, M. F. (November, 2013). Ignoring individual differences in belief bias leads to incorrect conclusions. Poster to be presented at the Psychonomic Society Annual Conference, Toronto, ON, Canada

Trippas, D., Handley, S. J., & Verde, M. F. (July, 2013). Individual differences in belief bias are important. Talk to be presented at the 7th annual London Reasoning Workshop at Birkbeck College

Trippas, D., Handley, S. J., & Verde, M. F. (November, 2012). Forced Choice Reasoning Eliminates Belief Bias. Poster presented at the Psychonomic Society Annual Conference, Minneapolis, MN, USA

Trippas, D., Handley, S. J., & Verde, M. F. (July, 2012). Forced Choice Reasoning Eliminates Belief Bias. Talk presented at the EPS Meeting at Bristol University.

Trippas, D., Handley, S. J., & Verde, M. F. (July, 2012). Smart People Show More Belief Bias. Talk presented at the 7th International Conference on Thinking at Birkbeck University, London

Trippas, D., Handley, S. J., & Verde, M. F. (September, 2011). Belief bias is not just response bias: A signal detection theory analysis. Talk presented at the 9th annual BPS Cognitive Section conference at Keele University.

Word count for the main body of the dissertation: 48,491

Signed:

22nd July, 2013

Chapter 1: Literature Review

1.1 Introduction

Humans reason to make sense of their environment. Doing so involves the application of logical rules of variable complexity. Reasoning is required for learning (Mitchell, De Houwer, & Lovibond, 2009), related to memory (Heit & Hayes, 2011), necessary for argumentation (Mercier & Sperber, 2011) and a major determinant of intelligence (Lohman & Lakin, 2009). It necessarily follows that no account of human cognition can be complete without an understanding of reasoning.

There are multiple types of reasoning, the two major ones of which are known as induction and deduction (e.g., Heit & Rotello, 2005; 2010). Induction involves the computation of beliefs through the generalisation from individual cases. For instance, if John has consistently woken up after going to sleep for the past 25 years, he will hold a strong belief that he will wake up the next time he goes to sleep. Deduction is concerned with the computation of truth based on what we know (or assume) is true. For example, if you assume that all humans are mortal and that John is human, then it is true that John is mortal. This is a logically valid inference. Deductive inferences are valid if the conclusion necessarily follows from the premises according to the classical rules of logic. Valid conclusions are true if the premises they are drawn from are true.

Rationality has been defined as the ability to conform to these logical rules (Stanovich, 1999). Piaget (1953), for instance, argued that cognitive development from irrational child to rational adult involved the acquisition of logical rules (Stenning & van Lambalgen, 2008). The study of deductive reasoning revealed that, in fact, people erred consistently against many of these logical rules (Wason, 1968). Because these structural deviations from the logical norm were considered irrational, they were termed biases. This fit in with the heuristics and biases approach in judgement and decision making popularised by

Tversky and Kahneman (1974). One example of such a bias in the domain of deduction was studied by Henle and Michael (1956). Their research showed that people who were instructed to judge the logical validity of certain types of deductive reasoning problems ignored logical rules. Instead, they based their judgements of validity on their prior beliefs (i.e., their knowledge of what is true in the world). This influence of prior beliefs on deductive reasoning is known as belief bias (Wilkins, 1928).

Belief bias was considered a particularly interesting example of irrationality in reasoning because it suggested that people resort to induction in tasks that require deduction. As such, belief bias was extensively studied as a paradigmatic example of irrationality in reasoning, although other systematic biases in deduction also occur. Evans, Barston, & Pollard (1983) investigated belief bias in three experiments which controlled for the impact of these various other reasoning biases. A large effect of validity was found, indicating that people could reason deductively – to a certain extent. Two reliable belief bias effects were also found. First, people accepted arguments with believable conclusions more than arguments with unbelievable conclusions, showing a first component to belief bias known as response bias. Second, logical validity and conclusion believability interacted, showing that the aforementioned validity effect appeared to be larger for unbelievable arguments. This logic by belief interaction was interpreted in terms of people engaging in more effortful reasoning for unbelievable arguments leading to increased accuracy. This type of belief bias is known as motivated reasoning (Lord, Ross, & Lepper, 1979).

Many studies have since replicated these two types of belief bias, i.e. response bias and motivated reasoning (e.g., Evans, Newstead, Allen, & Pollard, 1994; Evans, Handley, & Harper, 2001; Morley, Evans, & Handley, 2004; Newstead, Pollard, Evans, & Allen, 1992; Oakhill, Johnson-Laird, & Garnham, 1989; Klauer, Musch, & Naumer, 2000; Quayle & Ball, 2000; Torrens, Thompson, & Cramer, 1999). These studies resulted in the development of numerous belief bias theories. While many of these theories differ in their interpretation and explanation of the two belief effects, they all provide plausible explanations for the two key phenomena: the effect of prior beliefs on responding and the effect of prior beliefs on reasoning.

In a recent paper, Dube, Rotello, & Heit (2010) have challenged the motivated reasoning aspect of belief bias. Dube and colleagues analysed belief bias data using signal detection theory (SDT). They argued that a methodological oversight based on an incorrect assumption in all the previous work on belief bias led to an incorrect interpretation of the logic x belief interaction. On the basis of three experiments they argued that belief bias is just a response bias. In other words, beliefs only affect the response stage, not the reasoning stage. If their assertion holds true, all belief bias theories that provide a psychological explanation to account for the logic by belief interaction are incorrect, and much what has been assumed to be true about human reasoning is false.

In the remainder of this chapter we first introduce deductive reasoning in more detail by explaining logical validity, syllogisms, the link between rationality and bias, and a number of psychological theories which can account for the empirical findings. Next, we zoom in on belief bias by introducing the most viable belief bias theories and the methods which have been used to investigate it. Finally, a very recent theory of belief bias known as the response bias account is introduced. Given that the justification for the account is firmly rooted in SDT, we provide a detailed explanation of SDT after which we explain how it has resulted in a theoretical and empirical debate on the status of motivated reasoning in belief bias.

1.2 Deductive reasoning

Deductive reasoning is an important field of study. It has been argued that without deduction, science, technology, laws, social conventions, and culture would not exist

(Johnson-Laird & Byrne, 1991). Deduction differs from other kinds of reasoning (e.g., induction) in that it is primarily concerned with truth preservation: if the premises of an argument are true and the conclusion necessarily follows from these premises, then the conclusion is true. However, truth is not identical to logical validity: all true arguments are valid, but not all valid arguments are true. Arguments which are valid and which have true premises are known as sound arguments (e.g., Thompson, 1996). Much of the research on deductive reasoning has used unsound arguments as the general focus of interest lies with whether people can draw logically valid inferences.

1.2.1 Logical Validity and Syllogisms

An argument is logically valid if its conclusion *necessarily* follows from its premises (Johnson-Laird & Byrne, 1991). For instance, given the premises that "all sentences contain words" and that "all words contain letters", the conclusion that "all sentences contain letters" is logically valid because it necessarily follows. A potential alternative conclusion that "all letters contain sentences" is invalid given these premises, because it cannot follow from them. It is important to note that the semantic content of such a reasoning problem is independent of its logical validity. For instance, taking the previous valid argument and switching around the terms "sentences" and "letters" would yield the following equally valid argument: "all letters contain words", "all words contain sentences", "therefore, all letters contain sentences". These types of arguments are known as categorical syllogisms (hereafter, syllogisms).

Syllogisms are deductive arguments originally invented by Aristotle in an attempt to devise a formal system of reasoning. A syllogism is a reasoning problem which consists of three quantified sentences, two of which are premises and one of which is a conclusion. The conclusion is a statement about two end-terms (A and C) which are connected in the premises by a middle term (B). Every statement is preceded by one of the following quantifiers: All (universal affirmative, abbreviated A), No (universal negative, E), Some (particular affirmative, I), and Some … not (particular negative, O). The combination of the three quantifiers in a syllogism makes up the mood of the syllogism. For example:

Premise 1: Some artists are beekeepers

Premise 2: All beekeepers are clowns

Conclusion: Some artists are clowns

The mood of this syllogism is IAI (Some – All – Some). The ordering of the end- and middle terms in the premises makes up the figure of the syllogism. According to Aristotle's original notation the syllogism above (A – B; B – C) is in figure 1. In his book Prior Analytics, Aristotle originally only explicitly acknowledged three figures, but a fourth figure was introduced later. Confusingly, according to modern convention this form is known as figure 2, whereas Aristotle's original figures 2 and 3 are now respectively known as figures 3 and 4 (see Table 1.1). Finally, the conclusion direction of a syllogism can also be varied: it can go from A – C or from C – A.

Table 1.1

Syllogistic Figure: Old versus New Notation.

	Figure 1	**Figure 2**	**Figure 3**	**Figure 4**
Aristotle's convention	A – B B – C	A – B C – B	B – A B – C	/
Contemporary convention	A – B B – C	B – A C – B	A – B C – B	B – A B – C

In the remainder of this thesis the modern convention is used (Johnson-Laird & Byrne, 1991). Combining the 64 moods with the four figures and the two conclusion directions gives a total of 512 syllogisms, of which only 27 are logically valid.

1.2.2 Rationality and Bias

Originally, syllogisms were of particular theoretical interest because early scholars with an interest in human rationality assumed that rationality entailed the capability to reason logically, in other words, to think according to normative logical rules. In early research, syllogisms were used as a convenient benchmark against which reasoning performance could be evaluated, because their logical validity is clearly defined (Henle, 1962). In certain experiments, researchers presented people who were unfamiliar with syllogisms with a number of syllogistic premises and instructed them to come up with a valid conclusion. This is known as the conclusion generation paradigm (e.g., Markovits & Nantel, 1989). Others presented people with entire syllogisms, including the conclusion, and asked them to judge whether the conclusion was valid or invalid. This is known as the conclusion evaluation paradigm (e.g., Evans et al, 1983). It soon became apparent that regardless of the method used, people did not respond normatively. This caused researchers to conclude that people were irrational and that human reasoning was flawed (e.g., Stanovich, 1999; 2010). Interestingly, many people made the same mistakes in their reasoning. These systematic deviations from the supposed norm were called reasoning biases. Even though the idea that people's reasoning is irrational has been abandoned by most researchers (Elqayam & Evans, 2011; Gigerenzer & Brighton, 2009; Oaksford & Chater, 2007; 2009; although see Stanovich, 2010), these so-called biases became an interesting topic of study on their own because they can provide insights into the cognitive processes behind reasoning (Byrne & Johnson-Laird, 1990). We now introduce the four reasoning biases which all theories of syllogistic reasoning need to be able to account for (Khemlani & Johnson-Laird, 2012).

The atmosphere effect is the earliest explanation of how syllogistic reasoning is flawed (Sells, 1936; Woodworth & Sells, 1935). Participants were instructed to judge the validity of a number of syllogistic conclusions. The data revealed that the reasoners did not strictly adhere to the rules of logic as evidenced by their suboptimal reasoning performance. More importantly, many participants erred in very similar ways suggesting that they were influenced by the mood of the premises: when confronted with negative premises, reasoners preferred negative conclusions. If the premises contained the some-quantifier, reasoners preferred conclusions also containing the some-quantifier. Finally, if neither of these was the case, participants were biased towards accepting conclusions with the all-quantifier (Begg & Denny, 1969).

The concept of *illicit conversion* was formally studied by Chapman and Chapman (1959; see also Wilkins, 1928). They re-examined the atmosphere effect by replacing the conclusion evaluation paradigm (Sells, 1936; Woodsworth & Sells, 1935) with a multiple choice paradigm. Instead of a valid/invalid judgement, participants were presented with four possible conclusions and a fifth option that "nothing follows". Inspection of the error pattern revealed that the atmosphere effect could not completely account for the results. Instead, the researchers proposed that participants illicitly converted "All A are B" into "All B are A" and "Some A are not B" into "Some B are not A". By treating these premises as fully interchangeable, the participants generated an error pattern not explained by atmosphere alone. Whereas such conversions are appropriate for "Some A are B" and "No A are B", and even though they may often hold true in daily life, they are logically incorrect. A more formal model of illicit conversion has been introduced by Revlis (1975) and tested by Revlin, Leirer, Yopp, and Yopp (1980).

The figure effect is another common bias in syllogistic reasoning (Johnson-Laird, 1975; Wason & Johnson-Laird, 1972). Participants were presented with two logically equivalent pairs of premises, either in figure 1 (Some A are B, All B are C) or figure 2 (All B are A, Some C are B). For both pairs of premises they were given the choice between two valid conclusions: "Some A are C" or "Some C are A". Participants strongly preferred conclusions in the A-C direction for figure 1 and conclusions in the C-A direction for figure 2. Figural bias was less pronounced in figure 3 and completely absent in figure 4 (see Table 1.1 for an overview of the figures). These findings were later confirmed in experiments using non-abstract problem content and a conclusion generation paradigm (Johnson-Laird & Steedman, 1978).

Matching bias is a well-known phenomenon in several deduction tasks (Evans, 1972; Evans & Lynch, 1973). Applied to syllogistic reasoning, it makes the same predictions as the atmosphere effect, but for more sound theoretical reasons (Wetherick & Gilhooly, 1995). According to Wetherick and Gilhooly, qualitatively distinct subgroups of reasoners exist in any syllogistic reasoning task. One group attempts to apply logic to solve the syllogisms and manages to do so relatively successfully. A second group also attempts to apply logic, but is not successful in doing so. A final group of participants does not attempt to apply logic, but instead uses a simple matching heuristic to reach an answer. According to the matching hypothesis, these participants only endorse conclusions that are in the same mood as the most conservative quantifier. The "No"-quantifier is the most conservative one and the "All"-quantifier is the least conservative one. The "Some"- and "Some…not"-quantifiers are equally conservative and lie in between the other two. As such, both premises need be affixed by "All" in order for a conclusion using the "All" quantifier to be endorsed, whereas a single "No"-quantifier in either of the premises leads to the preference of a "No"-quantified conclusion. Importantly, this research introduced the powerful idea of different subgroups of reasoners doing qualitatively different things, a concept that will be explored further on in this dissertation.

1.2.3 Syllogistic Reasoning Theories

Syllogisms are often used as a means of investigating how people draw inferences. Given the heavy focus on this argument type, a viable theory explaining how people reason about syllogisms is required. Such a theory should be able to account for all the reasoning biases introduced in the previous section in a psychologically plausible way. We now introduce the four major classes of syllogistic reasoning theory: mental logic, mental models, verbal reasoning, and probability heuristics.

Mental logic is the idea that deduction through the use of logic is a fundamental human ability hard-wired into the brain (Rips, 1994; Braine & O'Brien, 1991; 1998; Braine & Rumain, 1983). According to mental logic reasoners engage in mental proofs using the rules of logic in order to determine the logical validity of deductive arguments such as syllogisms. This idea is rooted in the developmental theory of Piaget (1953) which assumed that as children grow up, they become more rational because they acquire an understanding of the relevant logical rules. One strong point of the mental logic tradition is the development of a computational model (PSYCOP) which is able to reason with syllogisms using predicate logic (Rips, 1994). Unfortunately, the model cannot convincingly account for the structural reasoning biases introduced above, limiting its usefulness. In order to solve this caveat, more recent iterations of the mental logic theory have become less rigorous in their assumptions, allowing for instance the inclusion of reasoning schemas which are influenced by prior knowledge (Braine & O'Brien, 1998). Currently, the idea that people possess a built-in mental system for deduction which conforms to the normative rules of logic has mostly been discredited (Evans, 2002; Elqayam & Evans, 2011).

Mental models theory (MMT) breaks with the mental logic tradition by arguing that people do not rely on built-in logical rules to reason. According to MMT, people construct mental models of the premises which they then use to deduce valid conclusions (Johnson-Laird, 1983; Johnson-Laird & Bara, 1984). These so-called mental models consist of mental tokens which represent the categories in the reasoning problem. Note that MMT is agnostic on the nature of these mental tokens (e.g., whether they are visual or symbolic). MMT posits a three-step process to explain how people reason about syllogisms. First, reasoners construct a model of the premises representing the current state of affairs. For instance, consider the premises “some artists are ballerinas” and “all ballerinas are cooks” the following mental model might be constructed:

artist	=	ballerina	=	cook
artist	=	ballerina	=	cook
		ballerina	=	cook

Second, reasoners attempt to come up with a true and informative conclusion that holds in the model of the premises which has been constructed so far. For instance, given the model of the premises above, the tentative conclusion that “all artists are cooks” might be drawn, because in this mental model all of the artists are also cooks. Third, when such a conclusion is found, reasoners will attempt to falsify it by constructing alternative models of the premises. The reasoner might construct for instance the following model of the premises as an alternative to the first one:

artist	=	ballerina	=	cook
artist	=	ballerina	=	cook
artist				

From this model, it can be seen that the initially drawn conclusion that “all artists are cooks” no longer holds, given the counterexample of the artist which is not a cook. Instead, the only valid conclusion given these two possible models of the premises is that “some artists are cooks”. In a conclusion evaluation paradigm where the participant has to judge the validity of a conclusion, a conclusion is considered invalid if a model which falsifies the conclusion can be found. If no such conclusion can be found, the reasoner considers it valid.

In contrast to mental logic, MMT can account for the syllogistic reasoning biases (Johnson-Laird & Bara, 1984). For instance, in the first stage of reasoning, figural bias drives the construction of the initial model, potentially affecting reasoning performance: if the reasoner does not search for counterexamples and if the initial model contains an invalid conclusion, it will be incorrectly accepted. One important aspect of MMT is that syllogisms vary in the amount of models that must be constructed to determine their validity. A syllogism such as “all artists are ballerinas”, “all ballerinas are cooks”, “therefore, all artists are cooks” is known as a one-model problem, because there is only one available (relevant) model of the premises:

artist	=	ballerina	=	cook
artist	=	ballerina	=	cook
artist	=	ballerina	=	cook

In the third stage, considering the generated counterexamples will be more difficult for syllogisms for which multiple models of the premises are available compared to ones that only allow for a single model of the premises, potentially due to the fact working memory constraints limit the amount of models which can be simultaneously considered (Johnson-Laird & Bara, 1984). This could then lead to structural errors that might be predicted by atmosphere, illicit conversion and matching (if for a different reason). One of MMT’s strengths lies with its ability to account for various types of deduction problems such as conditional inference, relational reasoning, and meta-deduction (Johnson-Laird & Byrne, 1991), making it a useful and plausible framework in the quest to account for human reasoning. MMT appears to be very good at explaining the syllogistic reasoning data patterns (Khemlani & Johnsson-Laird, 2012), but its account of conditionals is a lot less plausible in light of alternative ones such as for instance the suppositional theory of conditionals (Evans, Over, & Handley, 2005; although see also Barrouillet, Gauffroy, & Lecas, 2008).

The *verbal reasoning theory* (VRT) is a syllogistic reasoning theory also based on the mental models framework (Polk & Newell, 1995). According to the VRT, instead of searching for counterexamples, people reason by semantically encoding and re-encoding the premises. In line with the mental logic account’s PSYCOP, Polk and Newell wrote a number of computer programs of increasing complexity (VR1 – VR3) which generate putative conclusions based on an implementation of the linguistic encoding process which they propose underlies syllogistic deduction. The major difference between MMT and VRT is the absence of a falsification strategy rooted in the generation of counterexamples for the latter. Research, however, indicates that participants can search for counterexamples and occasionally do so (e.g., Byrne & Johnson-Laird, 1990).

The *probability heuristics model* (PHM) of syllogistic reasoning breaks with all the theories above by replacing the framework of traditional logic concerned with the computation of validity, with a probabilistic framework rooted in normative Bayesianism concerned with the computation of probabilistic- or p-validity (Chater & Oaksford, 1999; Oaksford & Chater, 2007; 2009). An argument is p-valid if its conclusion is less uncertain than its premises. Within this alternative normative framework, PHM proposes five heuristics which can account for typical syllogistic reasoning data patterns.

The first three heuristics focus on the generation of conclusions. First, the min-heuristic states that reasoners will use the least informative quantifier of the premises as the conclusion quantifier when generating conclusions. Ranked from most to least informative we have All > Some > No > Some…not. Second, according to the p-entailment heuristic, the second most preferred conclusion quantifier is one which is probabilistically entailed in the original conclusion. For instance “no dogs are animals” probabilistically entails that “some dogs are not animals”. Third, the attachment-heuristic states that reasoners determine the direction of the conclusion by looking at the subject of the premise with the least informative quantifier. If the subject of this so-called min-premise is an end-term (so either A or C), reasoners will use this term as the subject of the conclusion. If the subject of the min-premise is the middle term, reasoners will use the end-term of the max-premise as the subject of the conclusion. The final two heuristics are used by the reasoner to assess their confidence in the generated conclusions. According to the max-heuristic people are more confident in the generated conclusion if the quantifier of the max-premise (i.e., the premise with the most informative quantifier) is more informative. For instance, if the max-premise uses the all-quantifier as opposed to the no-quantifier, people will be more confident in the conclusion. Finally, the some…not-heuristic states that people will avoid producing or accepting some…not conclusions because they are so uninformative.

The p-validity framework in which the PHM is rooted shares some strengths with MMT. For one, the concept of p-validity can also provide plausible explanations for other types of tasks typically studied in the deduction paradigm, such as conditional reasoning. The major difference is that according to the PHM, participants engage in probabilistic reasoning rather than deduction (although there is some mention about probabilistic deduction in the new paradigm of reasoning, e.g., Evans & Over, 2012). The fact that the PHM also applies to statistical syllogisms using non-traditional quantifiers such as "Most" and "Few" is also a useful addition. Furthermore, PHM salvages the idea of human rationality by substituting the normative framework, suggesting that participants operate probabilistically. A negative point of PHM is the fact that according to the some…not-heuristic, participants should completely avoid accepting some…not conclusions due to their low informativeness. Yet, many of the most important experiments on syllogistic reasoning (e.g., Evans et al., 1983, Newstead et al., 1992) show that people do manage to reason about these conclusions relatively successfully. The concept of studying thinking and reasoning in function of p-validity substituted for traditional validity is gaining track within the new paradigm of reasoning (Evans, 2012). For the large majority of the traditional syllogisms, however, p-validity and validity are almost interchangeable, suggesting that the conclusions drawn from this dissertation can be interpreted in light of both the old and the new paradigms of reasoning.

1.3 Belief Bias

Belief bias is the tendency for people's deductive reasoning to be influenced by their prior beliefs in the truth of the conclusion. Consider the following syllogisms (Table 1.2) as a classic example from Evans et al. (1983):

Table 1.2

Four validity by believability syllogism types

	Believable	**Unbelievable**
Valid	Some addictive things are inexpensive No cigarettes are inexpensive ——— Some addictive things are not cigarettes 89%	Some millionaires are hard workers No rich people are hard workers ——— Some millionaires are not rich people 56%
Invalid	No addictive things are inexpensive Some cigarettes are inexpensive ——— Some addictive things are not cigarettes 71%	No millionaires are hard workers Some rich people are hard workers ——— Some millionaires are not rich people 10%

Note. Proportions indicate endorsement rates of conclusions as found by Evans et al. (1983)

Evans et al. (1983) instructed participants to judge whether the conclusions of syllogisms like these necessarily followed or not (i.e., whether they were logically valid). Even though for each validity level the arguments were identical in structure (valid: EI3_O1, invalid: IE3_O1), the believable conclusions were accepted much more than the unbelievable conclusions (47% more). The participants' prior beliefs in the truth of the conclusion caused their acceptance rates of these structurally identical arguments to vary wildly. One particular finding of interest in this case is the fact that for unbelievable arguments, the valid – invalid endorsement rate difference was much larger (46%) than for the believable arguments (18%).

Much like argument figure, for instance, conclusion believability affected the endorsement rates in syllogistic reasoning (Johnsson-Laird & Steedman, 1978). Contrary to the structural reasoning biases introduced in the previous section, however, belief bias can be seen as induction influencing deductive decisions, rendering it an interesting special bias worthy of additional study. The influence of prior beliefs on logical reasoning was well documented prior to Evans et al. (1983) (e.g. Wilkins, 1928; Henle, 1962; Henle & Michael, 1956). However, the experiments presented by Evans and colleagues were the first ones in which all the structural biases introduced earlier (i.e., atmosphere, figure, illicit conversion and matching) were adequately controlled for. Furthermore, logical validity and conclusion believability were crossed, rendering a two by two design with four cells: valid-believable, valid-unbelievable, invalid-believable and invalid-unbelievable (Table 1.2). This design allowed for a controlled analysis of the effects of logic and belief on the endorsement rates of conclusions. The majority of the studies on syllogistic belief bias published in the past 30 years which have used this methodology reliably found three statistical effects: a main effect of logic, a main effect of belief, and a logic x belief interaction effect. Given the reliability of these effects, all viable theories of belief bias must account both for the main effect of belief (i.e., why are people unable to suppress the influence of prior beliefs in the response stage?), as well as the interaction (i.e., why are people better at discriminating valid from invalid arguments when they have unbelievable conclusions?) as demonstrated by Evans et al. (1983).

With regards to the latter, some theories introduce the concept of motivated reasoning. Motivated reasoning entails that, rather than simply rejecting conclusions which go against their prior beliefs, people will engage in more effortful reasoning in order to try and disconfirm the conclusion. A conceptually similar example of motivated reasoning was discovered by Lord et al. (1979). They asked people to judge the methodological quality of various pieces of research. The results showed that people uncritically accepted the evidence of research studies whose conclusions were in line with their beliefs. Conversely, people were much more critical of the methodology of studies whose conclusions clashed with their prior beliefs, even though the arguments were structurally identical. For a syllogistic belief bias theory to be viable, it must explain both components, i.e. response bias and motivated reasoning. In the following paragraph the six traditional belief bias theories are introduced.

1.3.1 Belief Bias Theories

Selective Scrutiny (SS) is the first account of belief bias presented by Evans et al. (1983). According to SS, reasoners will uncritically accept believable conclusions, but will engage in more logical reasoning for unbelievable conclusions. This can explain both the main effect of beliefs and the logic x belief interaction. Counterevidence for selective processing comes from behavioural studies, many of which have shown that logic still has a significant effect on arguments with believable conclusions (e.g., Klauer, Musch, & Naumer, 2000). Logic should not affect believable arguments if participants accept all believable items without further scrutiny. Finally, response time studies (Thompson, Striemer, Reikoff, Gunter, & Campbell, 2003) have found that participants actually spend more time deliberating about believable than unbelievable conclusions, which is potentially incompatible with the idea that participants engage in more reasoning for unbelievable arguments.

Misinterpreted Necessity (MN) is an alternative theory of belief bias forwarded by Evans et al. (1983; see also Dickstein, 1980, 1981; Markovits & Nantel, 1989; Newstead et al., 1992). According to this account, participants have trouble understanding the concept of necessity required to adequately assess logical validity (i.e., that the conclusion has to follow in all cases for it to be logically valid). For certain types of invalid arguments (specifically known as indeterministically invalid) the conclusion can follow in some cases, but it does not follow in all cases. When people are confronted with this type of invalid argument during reasoning, they are thought to respond as a function of their beliefs: the fact that the conclusion can be both consistent and inconsistent with the premises is confusing and may be taken as evidence suggestive of validity. For valid arguments, the conclusion is always consistent with the premises, so beliefs will not be used as a response in this case. MN can account for both the main effect of belief and the logic x belief interaction. Note that the MN interpretation of the logic x belief interaction is not one of motivated reasoning, but rather one in terms of a response bias for invalid conclusions compared to an absence of response bias for valid arguments. The predictions from MN have not held up to scrutiny because it struggles to explain the belief effect on valid items. Furthermore, at least one study has found the logic x belief interaction for simple problems in which the conclusion is always consistent or inconsistent with the premises (Glinsky & Judd, 1994).

Metacognitive Uncertainty (MU) is an updated version of misinterpreted necessity which adds an overall response bias, thus allowing for a belief effect on valid arguments in addition to invalid arguments (Quayle & Ball, 2000). According MU, indeterminately invalid arguments require more effort to evaluate because all the alternative cases need to be kept in working memory simultaneously (i.e., the cases where the conclusion is consistent with the premises as well as the cases where the conclusion is not consistent with the premises). Participants for who the working memory resources required for this operation are exceeded will respond on the basis of believability for the invalid arguments, leading to a larger response bias on invalid than on valid items and as such the logic x belief interaction. Quayle & Ball (2000) demonstrated that participants with adequate working memory capacity did not show the logic x belief interaction, suggesting that they were able to complete this operation without an increased belief bias for invalid compared to valid problems. As with Misinterpreted Necessity, the interaction is not theoretically interpreted as originating from motivated reasoning. Contrary to MU's predictions, eye-tracking data suggests that participants do not spend more time generating alternative cases for invalid arguments (Ball et al., 2006).

Modified Verbal Reasoning Theory (MVRT) is a belief bias extension of the VRT introduced earlier (Thompson et al., 2003). MVRT is an attempt to reconcile the finding of a logic x belief interaction with the fact that participants actually spend more time on believable than on unbelievable conclusions (contrary to what a motivated reasoning accounts such as, for example, Selective Scrutiny would predict). According to this theory participants reason by semantically encoding and re-encoding the premises and will continue to do so until an adequate response is found or until a self-imposed deadline elapses. If they come up with a compatible response before the deadline, it will be accepted. If the deadline elapses, participants will either reject the conclusion or base their response on conclusion believability. Finally, MVRT argues that participants set an extended response deadline for believable conclusions because they are more palatable, resulting in a drawn-out rationalisation process. This theory is compatible with a traditional logic x belief interaction and longer response times for believable than unbelievable arguments. As MN and MU before, however, it does not interpret the logic x belief interaction in terms of motivated reasoning.

Mental Models Theory (MMT) has been adjusted and extended to account for belief bias in an evaluation paradigm by Oakhill, Johnson-Laird, & Garnham (1989). According to this version of MMT, participants will first construct a mental model of the premises from which they draw an initial conclusion. If the presented conclusion is inconsistent with the mental model, it is rejected. If the conclusion is consistent with the initial model, participants will assess the conclusion's believability. Believable conclusions are readily accepted. Unbelievable conclusions, on the other hand, will trigger a motivated reasoning process which entails the construction of alternative models. Only if the conclusion is consistent with all additionally constructed models, the conclusion is accepted. If a counterexample is found, the conclusion is rejected. For one-model problems, this leads to fairly good reasoning given that the initially constructed model will provide insight to the validity of the conclusion. Nevertheless, a small main effect of beliefs is still predicted due to an overarching response bias, here referred to as a "conclusion filtering mechanism" which takes place after the reasoning process. For multiple model syllogisms, MMT predicts the interaction between logic and belief because the additional search for counterexamples cued for unbelievable problems will lead to an increased probability of finding a disconfirmatory model which will lead to the increased rejecting of invalid problems. Newstead et al. (1992) found evidence in favour of MMT, with the absence of the logic x belief interaction for one-model problems. Counterevidence for MMT has been presented by Glinsky & Judd (1994), who found the logic x belief interaction for one-model problems. Ball, Phillips, Wade, and Quale (2006) used an eye-tracking procedure to investigate two predictions drawn from MMT. First, participants should look at unbelievable conclusions more than believable conclusions, because the latter are readily accepted. Second, participants should look at the premises longer if they first looked at an unbelievable conclusion. Their findings were incompatible with these predictions. A similar finding that participants take less time to respond to unbelievable than to believable conclusions provides additional evidence against MMT (Thompson et al., 2003; Thompson, Newstead, & Morley, 2011). The biggest potential issue for MMT, however, is the fact that participants typically only ever construct a single model of the premises (Evans, Handley, Harper & Johnson-Laird, 1999).

Selective Processing Theory (SPT) provides an alternative take on the mental models account and was independently proposed by Klauer et al. (2000) and Evans et al. (2001). According to SPT, participants only ever construct a single model of the premises (Evans et al., 1999) and reason from the conclusion to the premises (Morley et al., 2004). First, people assess the believability of the conclusion. If the conclusion is believable they will attempt to construct a model that is consistent with the premises. Given that for both valid and invalid multiple model problems such a model can be found, believable conclusions are generally accepted. If the conclusion is unbelievable people will attempt to construct a model that is inconsistent with the premises. For valid arguments, such a model cannot be found, leading to acceptance. For invalid arguments, an inconsistent model can be found, leading to an increased rejection rate. The net effect is increased logical performance for unbelievable compared to believable arguments. As such, the logic x belief interaction can be interpreted in terms of motivated reasoning (Evans, 2007). Note that SPT also predicts a general belief bias in terms of a response bias to operate for one-model and multiple-model problems. As with MMT, the finding that participants take longer to respond to believable arguments might be difficult to reconcile with the idea of increased reasoning for unbelievable arguments (Ball et al., 2006; Thompson et al., 2003; 2011).

1.3.2 Methods of Studying Belief Bias

All of the major belief bias theories have been developed in order to explain behavioural reasoning data. As we briefly explained above, there are multiple ways to collect belief bias data. In the conclusions generation paradigm, participants are presented with the premises of a syllogism and instructed to generate a conclusion (e.g., Markovits & Nantel, 1989). An advantage of this approach is that it resembles day-to-day reasoning: participants are given some information on the basis of which they have to draw a valid conclusion. A disadvantage of this approach is that participants do not always realise that the conclusion needs to contain both end-terms and that only the four syllogistic quantifiers can be used to reach a valid argument. This caveat can be overcome by carefully instructing participants about the expected form of the conclusion, but doing so inevitably reduces the ecological validity of the experiment because participants might no longer draw the conclusions they would do naturally.

An alternative option, known as the multiple-choice paradigm (e.g., Chapman and Chapman, 1959), can deal with the aforementioned problem. In this method participants are presented with the premises of a syllogism and provided with multiple conclusions, plus the option that "nothing follows". An advantage of this alternative method is that it increases reasoning performance and consequently facilitates investigation of the reasoning process. A disadvantage is that it might also introduce the use of unwanted (i.e., non-believability related) heuristics in conclusion selection. For instance, the probability heuristics model dictates that participants would never select the "some…not"-quantified conclusion option due to its low informativeness (Chater & Oaksford, 1999). Disregarding these types of syllogisms is not an option, however, because they are crucial for belief bias research. First, these syllogisms are the most complex because multiple models of their premises exist. This increased complexity could provoke a higher reliance on prior beliefs, in contrast to syllogisms for which the underlying logic is transparently easy (e.g., "all artists are ballerinas", "all ballerinas are cooks", "therefore, all artists are cooks"). The higher complexity of the multiple model syllogisms renders them a suitable platform for the study of belief bias. Second, the invalid counterparts of these syllogisms are indeterminately invalid, meaning that they are invalid because the conclusion does not necessarily follow, although it follows for certain models of the premises. Many belief bias theories (e.g., Metacognitive Uncertainty, Modified Verbal Reasoning Theory, Mental Models Theory, and Selective Processing Theory) make specific predictions for these types of problems, such as motivated reasoning induced accuracy increases or increased response bias. Consequently, these syllogisms are well-suited for the comparison of belief bias theories. Finally, syllogisms with "some…not"-quantified conclusions allow atmosphere, illicit conversion, and figural effects to be experimentally controlled for, rendering them useful from a methodological design point of view (Begg & Denny, 1969; Dickstein, 1978; Revlin, Leirer, Yopp, & Yopp, 1980).

The issues outlined above have led most belief bias research conducted in the past 30 years to adopt the conclusion evaluation paradigm in which participants are presented with a syllogism consisting of two premises and a conclusion. They are then instructed to respond “valid” if the conclusion is logically valid or “invalid” if it is not. The proportion of “valid” responses, also known as the endorsement rate, is then analysed by calculating three indexes. The logic index is a contrast of the number of “valid” responses to valid arguments (hits) minus the “valid” responses to invalid arguments (false alarms). This is an index of logical reasoning competence, with higher numbers indicating higher logical reasoning performance. The belief index is a contrast of the number of “valid” responses to believable arguments minus the number of “valid” responses to unbelievable arguments. This is an index of belief bias as response bias. Finally, the interaction index is a contrast of the logic index in the unbelievable condition and the logic index in the believable condition:

$$\text{Interaction Index} = (H_{\text{unbelievable}} - F_{\text{unbelievable}}) - (H_{\text{believable}} - F_{\text{believable}})$$

This is a measure of the degree to which participants reason better for unbelievable compared to believable arguments, i.e., motivated reasoning. Note that in more recent research, the logic, belief, and interaction index are not always explicitly mentioned because they are formally equivalent to the main effect of logic, the main effect of belief, and the logic x belief interaction in ANOVA of endorsement rates (i.e., proportion of “valid”-responses), respectively. Using this method, mixed evidence has been found for each of the six traditional belief bias theories. The recent application of a novel analytic technique, however, has raised some questions with the plausibility of all the work that has been conducted on belief bias so far.

1.4 The Response Bias Account

The logic x belief interaction is a key empirical observation which has inspired much theorising about which account of belief bias is to be preferred (e.g., the absence of the logic x belief interaction for one-model problems has been taken as evidence in favour of MMT and SPT and against SS). A recent paper by Dube et al. (2010) put this method of theorising into question. Dube et al. conducted three syllogistic reasoning experiments and analysed them using a method from the signal detection theory (SDT) framework. They showed that the interpretation of the logic x interaction in terms of motivated reasoning tacitly requires the assumption that hits (saying valid to a valid argument) and false alarms (saying valid to an invalid argument) are linearly related as response bias increases. This is unknown to or unacknowledged by many belief bias researchers. Their results led them to conclude that the logic x belief interaction takes on a different interpretation when analysed using SDT. Their analysis indicated that, rather than being an indication of the degree of motivated reasoning, the interaction is merely a statistical artefact. Consequently, according to their findings, only response bias remains. If their claims hold, then all accounts of belief bias which have some psychological mechanism in place to account for the interaction (i.e., the six theories introduced in the previous section) are necessarily incorrect. Dube et al. propose an alternative account which argues that belief bias is just a response bias – the response bias account of belief bias. Before describing Dube et al.'s findings in more detail, we now turn to an explanation of SDT.

1.4.1 Signal Detection Theory

SDT originated in the field of radar research to deal with the fact that radars were limited in performance due to intrinsic noise and that any "blip" could originate not only from a mixture of signal and noise, but also from noise alone (Peterson, Birdsall, & Fox, 1954). Green and Swets (1966) realised that the same concepts applied in psychophysics, the study of how human perception is related to the physical world (Fechner, 1860). They adapted SDT for use with humans as the detecting agent. The main tenet of SDT is the idea that for any yes–no decision about the presence of a stimulus both targets (signals) and non-targets (noise) exist, and that they can both lead to positive (i.e., yes, the target was presented) and negative (no, the target was not presented) responses. Ever since its original introduction SDT has been applied to many fields in cognitive psychology, such as memory (Macmillan & Creelman, 2005), social cognition (Gable, Reis, & Downey, 2003), and reasoning (Heit & Rotello, 2005; 2010; Rotello & Heit, 2009).

The main assumption of SDT is that targets and non-targets are normally distributed on a strength dimension, with the targets usually having a higher mean than the non-targets (see Figure 1.1).

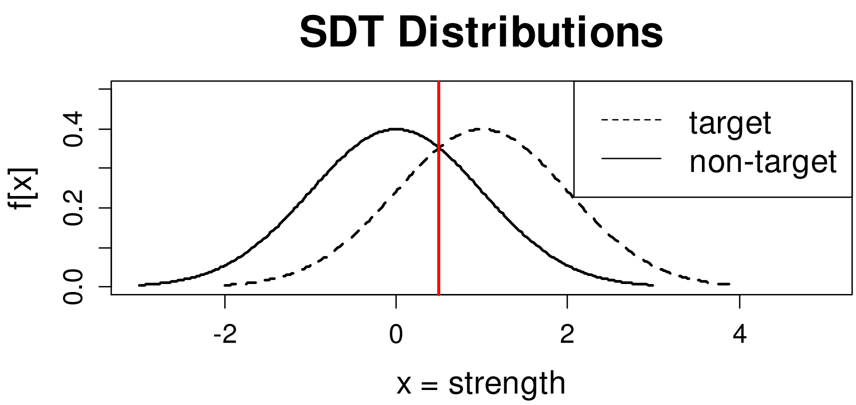


Figure 1.1 Example target (signal) and non-target (noise) distributions assumed to underlie human decision making performance by the standard SDT model. The vertical line is the response criterion.

In a typical SDT experiment on human perception, participants are presented with a number of stimuli varying in strength, for instance tones of variable loudness. They are then instructed to respond “yes” if they heard a tone and “no” if they did not. As an added difficulty, on certain trials no tone is presented. With loud tones the signal is easy to detect, but the less loud the tone, the more difficult it becomes to detect the target (i.e., to distinguish signal from noise). With this type of study design there are four possible response categories. Two of these are correct responses: the participant can respond “yes” to a tone that has been presented. This is known as a hit (H). The participant can also respond "no" if no tone has been presented, which is known as a correct rejection (CR). There are also two incorrect response classes: the participant can respond "yes" even if no tone has been presented. This is known as a false alarm (F). Finally, the participant can respond "no" even though a tone was presented, which is known as a miss (M), see Table 1.3.

Table 1.3

Four response classes according to SDT.

	Target presented	**Non-target presented**
"Yes" response	Hit (H)	False Alarm (F)
"No" response	Miss (M)	Correct Rejection (CR)

A large disadvantage of analysing the proportion of correct responses in a yes-no task without taking these four possible response categories into account is that an increase in the proportion of reported detections does not necessarily equate to an increase in detection performance. For instance, imagine a doctor who has to investigate X-rays in order to assess whether patients have a tumour or not. Perhaps he correctly classifies a lot of tumours because he is very good at detecting the blobs on the X-Rays. If this is the case, then most of the doctor's responses are hits and few are false alarms, resulting in high detection performance (*sensitivity* in SDT terms). It is also possible, however, that he is not very good at interpreting the X-rays, but that he simply responds "yes" quite often. If this is the case, even though he has a large proportion of hits, he also has a large amount of false alarms, resulting in poor sensitivity. In SDT terms, the tendency to be inclined to give more yes than no responses (or vice versa), regardless of the underlying stimuli is known as *response bias*. A marked preference of yes to no responses is known as liberal responding, whereas the opposite is termed conservative responding. Contrary to many traditional analysis techniques, SDT acknowledges that sensitivity and response bias can vary independently. Furthermore, the theory provides a framework which allows sensitivity to be disentangled from response bias.

Sensitivity is related to the relative positioning of the target and non-target distributions. A larger distance between both distributions results in greater sensitivity. The rightward shift of the target distribution in the bottom panel of Figure 1.2 compared to the target distribution in the top panel demonstrates this concept. Sensitivity (d') can be calculated using the values in Table 1.3 using the following formula (Macmillan & Creelman, 2005):

$$d' = z(H) - z(F)$$

(z(x) is a function which gives the inverse of the standard normal cumulative distribution)

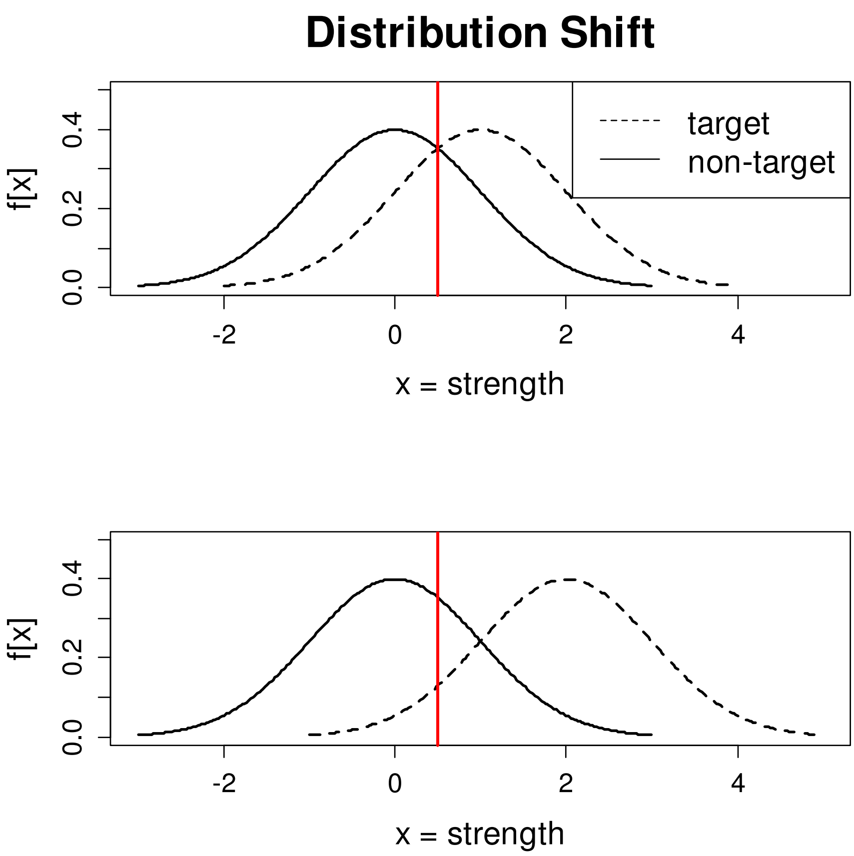


Figure 1.2 The target distribution of the bottom panel is shifted rightwards compared to the target distribution in the top panel, indicating increased sensitivity for the former.

Response bias is a measure of how likely the participant is to respond "yes" or "no". SDT acknowledges the fact that, all things being equal, certain people are more likely to respond yes (they are more liberal in their judgements) and that others are more likely to respond no (they are more conservative; Aminoff et al., 2012; Kantner & Lindsay, 2012). Likewise, certain conditions will render participants more liberal on average, whereas others might lead to more conservative judgements overall. Formally, the degree of response bias is equated to the position of the response criterion (c). If the strength of a stimulus (regardless of its target-status) exceeds the criterion, the participant will respond positively. The bottom panel of Figure 1.3 demonstrates how a downward criterion shift indicates more liberal responding. Bias can be calculated using the values in Table 1.3 with the following formula:

$$c = -0.5 * [z(H) + z(F)]$$

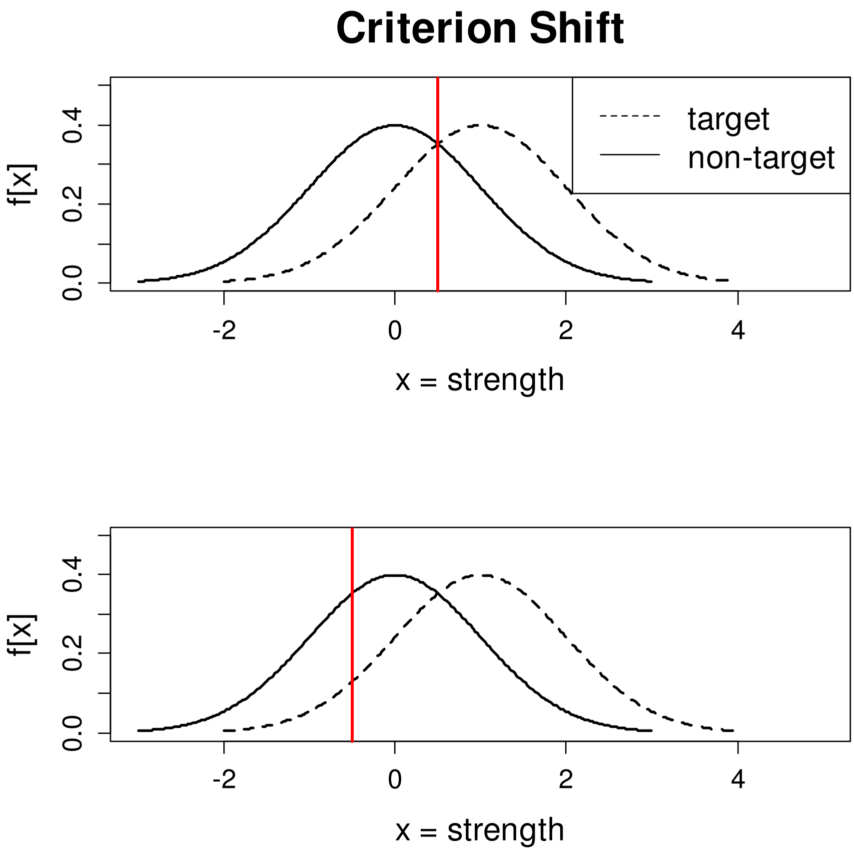


Figure 1.3 The response criterion (vertical line) in the bottom panel is shifted leftwards compared to the criterion in the top panel, indicating a more liberal response bias.

A Gaussian distribution is defined by two parameters: its mean (μ) and its standard deviation (σ). One problem of the single-point sensitivity (d') and bias (c) measures is that they require the assumption of equal variance (σ^2) in the target and non-target distributions. Studies show that, more often than not, this assumption does not hold (Green, 1986), for instance because the target distribution is more spread out than the non-target distribution. Fortunately, SDT can deal with the unequal-variance problem in multiple ways. One often used way is known as the receiver operating characteristic (ROC) curve method.

The ROC Method

A ROC curve is a cumulative plot of hits versus false alarms across increasingly liberal levels of responding. Otherwise put, every point on the ROC implies an identical sensitivity level, but points towards the right indicate a greater tendency to say yes (Macmillan & Creelman, 2005). An example of a ROC implied by a standard equal variance SDT model, the distributions of which are presented in Figure 1.1, can be found in Figure 1.4A. ROCs are useful tools for data interpretation because three crucial things can be derived from them. First, the distance of the ROC to the upper left corner indicates sensitivity (independent from bias). The closer the ROC is positioned to the upper left corner, the higher sensitivity. The ROC in Figure 1.4B shows higher sensitivity compared to the ROC in Figure 1.4A. Second, the relative shift of an ROC to the right indicates a more liberal response bias (for identical levels of sensitivity). The ROC in Figure 1.4C shows an example of more liberal response bias compared to the ROC in Figure 1.4A. Finally, the shape of the ROC provides some information about the shape of the underlying distributions. Curvilinear ROCs are consistent with (but do not necessitate) normally distributed targets and non-targets, as expected by SDT. The observation of a linear ROC would suggest the SDT model is inappropriate because the underlying distributions are not normal, see Figure 4D.

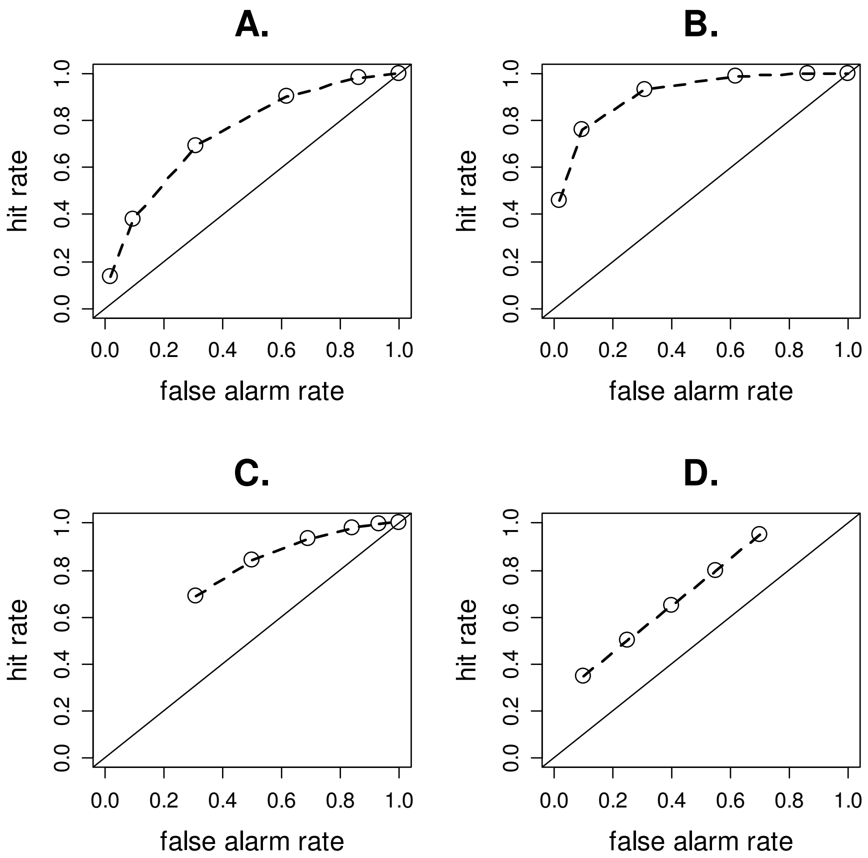


Figure 1.4 A. ROC implied by standard equal variance SDT model. B. ROC implied by model with higher sensitivity .C. ROC implied by model with equal sensitivity as in A. but with a more liberal response criterion. D. ROC implied by non-SDT model. The diagonal line where false alarm rate = hit rate shows chance performance.

Empirical ROCs can be constructed by instructing participants to adjust their response criterion while keeping sensitivity constant. The most common way to achieve this is by collecting confidence ratings (e.g., Yonelinas & Parks, 2007). After each yes-no response participants are asked to rate how confident they are in their response, for instance on a scale from 1 (not very confident) to 3 (very confident). The binary judgements are then combined with the confidence ratings to create a 6-point scale: 6 (high confidence yes response), 5 (medium confidence yes response), 4 (low confidence yes response), 3 (low confidence no response), 2 (medium confidence no response), 1 (high confidence no response). The leftmost point on the ROC is a plot of P(6 | target) v. P(6 | non-target). The next point along the curve is P(6 + 5 | target) v. P(6 + 5 | non-target), and so forth. Given the cumulative nature of these response categories, the final point of the ROC must necessarily fall at (1, 1). The major advantage of ROC curves over single-point measures (i.e., d' and c) is that the latter require the assumption of equal variance in the target and non-target distributions. ROCs do not require this assumption, and as such they can be used to calculate more appropriate sensitivity (e.g., d_a and A_z) and bias (e.g., c_a) parameters.

1.4.2 The Response Bias Account

Dube et al. (2010) applied the ROC procedure to a typical belief bias experiment. They instructed participants to reason about syllogisms for which validity was crossed by believability, resulting in the traditional item types (i.e., valid-believable, valid-unbelievable, invalid-believable, invalid-unbelievable, see Table 1.2). In addition to the standard conclusion evaluation paradigm, participants were asked to rate their confidence in each validity judgement. The data were analysed in two ways: the endorsement rates were analysed in the traditional manner using ANOVA and the ROCs were analysed by fitting and comparing SDT models. The endorsement rate analysis resulted in the standard three effects: a main effect of logic, a main effect of belief, and a logic x belief interaction, suggesting the typical belief bias effects were present (i.e., response bias and motivated reasoning). The ROC curves (Figure 1.5) painted a different picture. The first notable point was that the ROCs were curvilinear instead of linear. This was incompatible with the linear ROCs predicted by the logic x belief interaction (or interaction index). Some simple algebra demonstrates why the interaction index necessarily predicts linear ROCs. The interaction index [$II = (H_U - F_U) - (H_B - F_B)$] is a contrast of the logic index for unbelievable [$LI_U = H_U - F_U$] and believable [$LI_U = H_U - F_U$] problems. Solving both logic indexes for hits shows that in both cases, the hit rate is linearly related to the false alarm rate [$H_U = F_U + LI_U$; $H_B = F_B + LI_B$]. Consequently, the interaction index is only a valid measure of the difference in reasoning performance between the believable and unbelievable conditions if the ROCs are linear. In Figure 1.5 it can be seen that the hypothetical linear unbelievable (red) ROC is shifted towards the upper left compared to the linear believable (black) ROC, supporting the significant logic x belief interaction in the endorsement rate analysis. However, the empirical ROCs were curvilinear, demonstrating the inappropriateness of the interaction index (Dube et al., 2010).

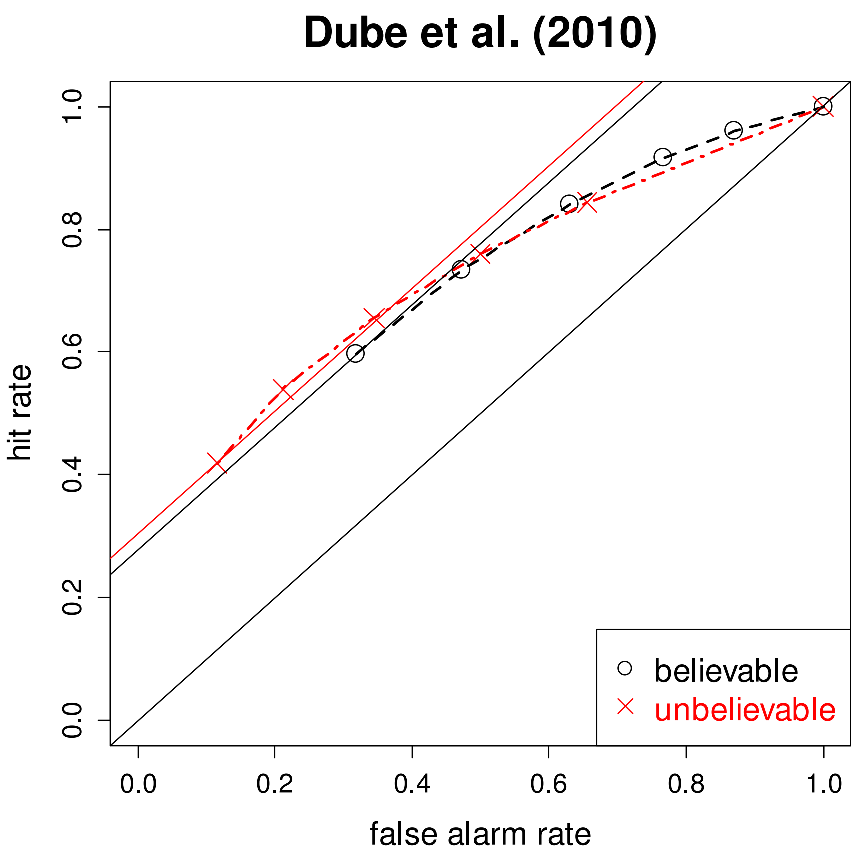


Figure 1.5. ROC curves presented by Dube et al. (2010, Experiment 2). Red and black lines are ROCs implied by the interaction index.

Having demonstrated that the experimental data is incompatible with the linearity assumption required for the correct interpretation of the logic x belief interaction, Dube et al. instead applied the more appropriate SDT procedure. Visual inspection of the ROCs suggested that the believable ROC was shifted to the right compared to the unbelievable ROC which is compatible with a response bias. Furthermore, the unbelievable ROC did not seem to be closer to the upper left than the believable ROC, suggesting that accuracy did not differ between the conditions. These findings were confirmed using formal model fits. Dube et al. used a maximum likelihood procedure to fit a standard unequal variance SDT (UVSD) model to the data. They then performed likelihood ratio tests to compare the unconstrained model with a model in which there was no accuracy difference between the believable and unbelievable condition. The test showed that constraining these parameters did not impede model fit. Constraining the response bias parameters to be equal across conditions, however, did significantly reduce the fit of the model. Taken together, these findings suggested that belief bias is just a response bias. The far reaching consequence of this is that all major theories of belief bias which explain the logic x belief interaction in terms of a psychological mechanism (motivated reasoning or otherwise) are incorrect (Dube et al., 2010).

1.4.3 The Belief Bias Debate

In response to Dube et al.'s conclusions that belief bias is just a response bias, Klauer and Kellen (2011) applied a different measurement method to assess the viability of this strong conclusion. They applied a multinomial processing tree (MPT) model to Dube et al.'s data and found a) that the MPT model fit the data better and b) that beliefs did have an impact on reasoning accuracy according to this alternative model, in line with the traditional theories of belief bias. Dube, Rotello, & Heit (2011) in turn replied to Klauer and Kellen by comparing the flexibility of the SDT and MPT models using model recovery simulations and demonstrated that the latter was more flexible, leading to its superior fit. They concluded that the SDT model was more appropriate and stuck with their original conclusion that belief bias is just response bias.

This theoretical debate and its heavy methodological focus have left the field of belief bias in a state of uncertainty with regards to the status of motivated reasoning. Depending on the measurement method used (SDT v. interaction index / MPT), belief bias is just a response bias or also contains a motivated reasoning component. Given the far-reaching conclusions of this research and the fact that it is based on the relatively small amounts of empirical data currently available, we set out to further the belief bias debate. More specifically, in this dissertation we aimed to investigate the status of motivated reasoning and response bias using SDT, arguably the most appropriate method (Dube et al., 2011). We now provide an outline of the structure of this dissertation.

1.5 Dissertation Structure

In this first chapter we have outlined why belief bias is an important but as of yet poorly understood topic of study within cognitive psychology. A debate in the literature about which aspects of belief bias exist and which methods are the most appropriate to investigate this problem is currently in progress (Dube et al., 2011; Klauer & Kellen, 2011). The aim of the research presented in this dissertation is to advance the debate between the traditional interpretation of belief bias and the alternative interpretation that belief bias is just a response bias. In the following chapters we present research designed specifically to further this debate using two methods within the SDT framework. In Chapter 2, a novel SDT-based method known as forced choice reasoning is introduced in an attempt to find converging evidence for the response bias account. In Chapter 3, we use the SDT-based ROC method to investigate the effect of various experimental manipulations on belief bias. In Chapter 4, both methods are used to investigate the impact of individual differences in cognitive ability and analytic cognitive style, in an attempt to reconcile the empirical differences observed. In Chapter 5 we discuss the theoretical viability of the major belief bias accounts in light of the evidence presented in the literature and this dissertation. We now turn to Chapter 2 in which we introduce the forced choice reasoning method.

Chapter 2: Forced Choice Reasoning

2.1 Introduction

In Chapter 1 we provided an overview of the methods typically used to study belief bias. Evidence suggested that the frequently used conclusion evaluation paradigm is suboptimal for the study of belief bias and that interpreting results from this method might lead to incorrect conclusions. Dube, Rotello, & Heit (2010) argued that the signal detection theory (SDT) framework is a more appropriate measurement method for the study of belief bias, given the observed ROC curvilinearity of reasoning data (Dube et al., 2010; Heit & Rotello, 2005; 2010; Rotello & Heit, 2009). Plotting ROCs and fitting SDT models to them is a better way of measuring belief bias, but alternative SDT-based approaches also exist. In this chapter, we introduce one such approach known as the forced choice method. The main aim of the current chapter is to test the viability of the response bias account of belief bias using this novel method. Converging evidence consistent with the response bias account would strengthen its theoretical position as the most viable explanation of belief bias. Evidence incompatible with the response bias account in the form of motivated reasoning, however, would cue the start of a more thorough search into the specific mechanisms underlying (both components of) belief bias.

2.1.1 Forced Choice Reasoning

As explained in Chapter 1, the ROC procedure provides a method which allows for the estimation of reasoning accuracy independently from response bias. Research on psychophysics has shown that this is not always the most efficient option to study a phenomenon (Fechner, 1860). In contrast detection research, for example, higher discrimination accuracy was attained by presenting two stimuli next to each other and by instructing the participant to choose the brighter one. This research paradigm is known as the forced choice method (Macmillan & Creelman, 2005). The forced choice procedure is an alternative method within the SDT framework which can achieve bias-free estimates of accuracy by controlling for response bias. The main difference between the ROC and the forced choice methods is that in the former only one stimulus is evaluated per trial, whereas in the latter, multiple stimuli have to be judged per trial. The forced choice method is often used in research on recollection memory, where the task is to study a list of words and to indicate in a test phase whether a word was studied (old) or not (new). Instead of presenting targets intermixed with lures and instructing the participants to judge whether each word is old or new, two words are presented next to each other. It is the participant's task to choose which of the two is old (Macmillan & Creelman, 2005).

One notable example of how the forced choice method is capable of increasing our understanding of a phenomenon in a way that the more traditional single presentation methods cannot is found in research on the frequency effect in recognition memory. The frequency effect entails that people are more likely to respond “old” to words which have a higher frequency in real life due to their increased familiarity. As such, in the single presentation paradigm, when participants responds “old” we have no way of knowing whether they did so because they remembered the word or because it has a high frequency. In the forced choice presentation method, however, the presented target and lure stimuli can be chosen to be of equal frequency. Consequently, frequency-based response bias is controlled for because the participants cannot choose one word over the other on the basis of its frequency. Instead, judgements mainly reflect the memory strength of the stimulus (Glanzer & Bowles, 1976). An additional motivated for using the forced choice method alongside the ROC method is that it does not require various modeling assumptions to be met. Occasionally, some scepticism or doubt as to whether the required assumptions for single-point or ROC measures are met exists. In these cases, converging evidence using an alternative method within the same framework can strengthen the viability of the results.

We applied the forced choice method to reasoning in order to study belief bias. The previously introduced methods for studying belief bias (i.e., the conclusion generation paradigm, the conclusion evaluation paradigm, the multiple choice method, and the ROC method) all have in common that only one argument is presented at a time. Participants have to evaluate on each trial whether the presented argument is valid or invalid. One major advantage of the forced choice method is that it can control for the unwanted effects of individual differences in response bias. Individual differences in criterion placement exist (Animoff et al., 2012; Kantner & Lindsay, 2012) and might introduce additional noise in the assessment of sensitivity (Benjamin, Diaz, & Wee, 2009).

We adapted the forced choice procedure and tailored it to be suitable for the study of belief bias in order to test Dube et al.'s (2010) claim that belief bias is just a response bias. To our knowledge, this is the first time the forced choice method has been applied to reasoning. In Experiment 1, we use the forced choice reasoning method to conduct a direct test of motivated reasoning independently of response bias. Positive evidence of motivated reasoning would suggest that the response bias only account is insufficient to explain belief bias.

2.2 Experiment 1

We extended the forced choice method for the study of belief bias as follows: on each trial two arguments are presented, one of which is valid and one of which is invalid. The participant has to evaluate both arguments and discriminate the valid argument from the invalid by reasoning about both. The major advantage of the forced choice approach applied to reasoning is that the possibility for belief bias as response bias can be eliminated. Participants know that one argument is valid and that one is invalid, but they also notice that both conclusions have the same believability status (e.g., in Table 2.1 both arguments have a believable conclusion). Consequently, a simple belief-based decision heuristic (i.e., accept believable, reject unbelievable) cannot be applied. Believability can be manipulated between trials such that on some trials both conclusions are believable and on other trials both conclusions are unbelievable. This novel forced choice paradigm allows for a direct test of motivated reasoning by comparing accuracy for the believable and the unbelievable problem types. Unlike most of the traditional accounts of belief bias (see Chapter 1), the response bias account predicts that beliefs do not influence reasoning accuracy (Dube et al., 2010). Therefore, an effect of beliefs on accuracy would be incompatible the response bias account. We also included a neutral condition to investigate whether unbelievable arguments increase reasoning performance, as suggested by motivated reasoning accounts (e.g., Evans et al., 2001), or whether believable arguments lead to worse reasoning.

Table 2.1

Example of Syllogisms Presented in a Typical Forced Choice Reasoning Trial.

VALID BELIEVABLE:	INVALID BELIEVABLE:
No cigarettes addictive things are inexpensive	No addictive things are inexpensive
Some addictive things are inexpensive	Some cigarettes are inexpensive
________________	________________
Some addictive things are not cigarettes	Some addictive things are not cigarettes

Note. Both syllogisms have identical conclusions. The left syllogisms is valid, the right one invalid.

2.2.1 Method

Participants

A total of 22 participants took part in exchange for course credit. Seven participants were male and 15 were female (age 18 – 31, M = 21, SD = 3).

Materials

A list containing 48 problems was created to be used as stimuli. Every problem consisted of two syllogisms presented side by side. The two syllogisms always had identical conclusions. One syllogism was always valid and the other one was always invalid. Half the time the syllogism on the left was valid and half the time the one on the right was valid. Sixteen problem frames were repeated three times (see Table 2.2). Half of these were valid and half were invalid. All syllogisms were of the multiple-model type, half of them two-model and half three-model. For the two-model syllogisms, logical validity was manipulated by changing the quantifier of one of the premises from “All” to “No”. For the three-model syllogisms, logical validity was manipulated by reversing the conclusion direction. A syllogism is defined as logically valid if its conclusion necessarily follows from the premises. If the conclusion is possible, but not necessitated by the premises, the argument is logically invalid.

Table 2.2

Syllogisms Used in Experiments 1-11.

Valid		Invalid	
EI3_O2:	No A are B	EI3_O1:	No A are B
	Some C are B		Some C are B
	Some C are not A		Some A are not C
EI4_O2:	No B are A	EI4_O1:	No B are A
	Some B are C		Some B are C
	Some C are not A		Some A are not C
EI1_O2:	No A are B	EI1_O1:	No A are B
	Some B are C		Some B are C
	Some C are not A		Some A are not C
IE1_O1:	Some A are B	IE1_O2:	Some A are B
	No B are C		No B are C
	Some A are not C		Some C are not A
OA3_O1:	Some A are not B	OE3_O1:	Some A are not B
	All C are B		No C are B
	Some A are not C		Some A are not C
AO3_O2:	All A are B	EO3_O2:	No A are B
	Some C are not B		Some C are not B
	Some C are not A		Some C are not A
OA4_O2:	Some B are not A	OE4_O2:	Some B are not A
	All B are C		No B are C
	Some C are not A		Some C are not A
AO4_O1:	All B are A	EO4_O1:	No B are A
	Some B are not C		Some B are not C
	Some A are not C		Some A are not C

Note. A = "All X are Y"; E = "No X are Y"; I = "Some X are Y"; O = "Some X are not Y". The first digit indicates the figure (1 = AB-BC; 2 = BA-CB; 3 = AB-CB; 4 = BA-BC). The second digit indicates the conclusion direction (1 = A-C; 2 = C-A).

Potential unwanted effects of problem content were controlled for by randomly assigning the problem contents to the syllogisms for each participant anew (e.g., Klauer & Singmann, in press). This resulted in a unique problem list for every participant. Conclusion believability was manipulated by combining sixteen superordinate categories (e.g., birds) with two subordinate members of each those categories (e.g., parrots and sparrows; see Table 2.3 for a full list), resulting in relationships that were either definitionally true or false.

Table 2.3

Item Contents Used in Experiments 1-11.

Category	**Members**			
amphibians	frogs	salamanders	toads	newts
birds	parrots	sparrows	ducks	robins
boats	kayaks	canoes	yachts	speedboats
cars	BMWs	Volvos	Vauxhalls	Fiats
criminals	robbers	murderers	embezzlers	terrorists
furniture	desks	sofas	cupboards	bookcases
dogs	Spaniels	Labradors	Terriers	Dalmatians
drinks	beers	sodas	wines	whiskeys
fish	trout	salmons	cods	haddocks
fruits	prunes	peaches	apples	bananas
insects	bees	beetles	ants	spiders
reptiles	lizards	iguanas	snakes	crocodiles
tools	hammers	saws	spanners	shovels
trees	oaks	willows	pines	maples
vegetables	carrots	cabbages	parsnips	radishes
weapons	cannons	swords	guns	spears

Note. In Experiment 1 only the first two columns of members were used for the believable and unbelievable problem types, whereas the third column was used as the linking term for the neutral problems.

An example of a believable conclusion is: “some birds are not parrots” (because prior knowledge dictates that there are other types of birds besides parrots). An example of an unbelievable conclusion is: “some sparrows are not birds” (because prior knowledge dictates that all sparrows are birds). Premise believability is known to exert a response bias in addition to the conclusion believability effect (Thompson, 1996). Given the main focus on conclusion believability, premise believability was controlled for. Sixty-four pseudowords (Table 2.4) were generated using Wuggy (Keuleers & Brysbaert, 2010) and

48 were randomly assigned to the middle terms of the syllogisms. Nonsense middle terms are commonly used to control for premise believability (e.g., Newstead, Pollard, Evans, & Allen, 1992; Evans, Handley, & Harper, 2001).

Table 2.4

Pseudowords Used in Experiments 1-11.

redes	fosks	pives	pields	decottions	sothods	renes	bunges
wasses	geets	swants	cronxes	firters	nickhomes	revoules	pinds
foins	chindles	soats	sonds	pumes	papes	trops	envenches
lebs	brops	stoges	crots	punties	stamuses	vennars	cortemns
weens	quinces	loaxes	stoals	curges	gruts	cosuors	wightes
punds	jubs	parfs	fises	hoons	tutches	brimbers	punes
cofts	spashes	fimps	brams	heets	piffures	burtes	queels
flamps	dathses	darms	vosts	trinnels	goples	boodings	veemers

Note. Pseudowords were used to control for premise believability.

Believable, unbelievable and neutral problems trials were created to be used as stimuli. Believable problem trials featured a valid-believable syllogism beside an invalid-believable one (see Table 2.1 for an example). Believable conclusions were created by presenting the superordinate category before a subordinate member of that category (e.g., some birds are not parrots). Unbelievable problem trials featured a valid-unbelievable syllogism next to an invalid-unbelievable one. Unbelievable conclusions were created by reversing the assignment order used in the believable case (e.g., some parrots are not birds). Neutral problem trials featured a valid-neutral syllogism next to an invalid-neutral one. Neutral conclusions were created by presenting two pseudowords (e.g., some fromps are not blarks). The middle-term of the premises of neutral problems was always assigned a randomly chosen subordinate member that was unused in the believable and unubelievable problems. Combining these syllogistic structures and item contents yielded a uniquely randomised list of 48 problems for every participant. Each list contained 16 believable problems, 16 unbelievable problems and 16 neutral problems.

Procedure

The participants were tested on individual computers in small groups. Upon entering the lab they were sat in front of a computer. They were presented with an information sheet outlining the general aims and procedure of the experiment, a form allowing them to give informed consent, and a detailed instruction sheet which explained the meaning of logical validity using a definition and an example (see Appendix A). The experimenter instructed the participants to subsequently read the information sheet, to sign the consent form if they agreed, and to read the instructions. Participants then had to enter their age and gender on the screen using the keyboard and mouse. Further instructions were presented on the screen using E-Prime 2.0. The instructions read:

"In the following experiment you will repeatedly be presented with two reasoning problems. One of these problems is logically valid. One of these problems is logically invalid. It is your task to choose which problem is valid. Use the mouse to click the box that contains the valid argument according to you. The box around the reasoning problem you chose will turn green, indicating that you think it is valid. The box around the reasoning you didn't choose will turn red, indicating that you think it is invalid. After every choice, use your mouse to indicate how confident you are that you made the correct decision. Before the actual experiment you will be presented with 6 practice trials to ensure that you understand the instructions. After the practice trials the actual experiment begins. If you have any questions ask the experimenter before or during the practice trials."

The participants then completed six practice trials (two believable, two unbelievable and two neutral). No accuracy feedback was provided. If the participants had no further questions after completing the practice trials they were told to complete the 48 experimental trials at their own pace. For every problem participants would be presented with two syllogisms side by side. Both syllogisms always had identical conclusions and item contents. Upon choosing one of the response options by clicking on it, the border of the box containing the chosen syllogism would turn green and the word "VALID" appeared underneath. Simultaneously the border of the box containing the remaining syllogism would turn red and the word "INVALID" appeared under it (see Figure 2.1 for an example). Participants were allowed to change their selection an unlimited amount of times by clicking the other box, but only chose to swap from one option to the other in 2% of the cases. The participants had to confirm their final choice by rating their confidence on a scale from 1 (not very confident) over 2 (moderately confident) to 3 (very confident) by clicking in the corresponding box, after which the next problem trial was presented. Upon completing all the experimental trials participants were thanked and debriefed.

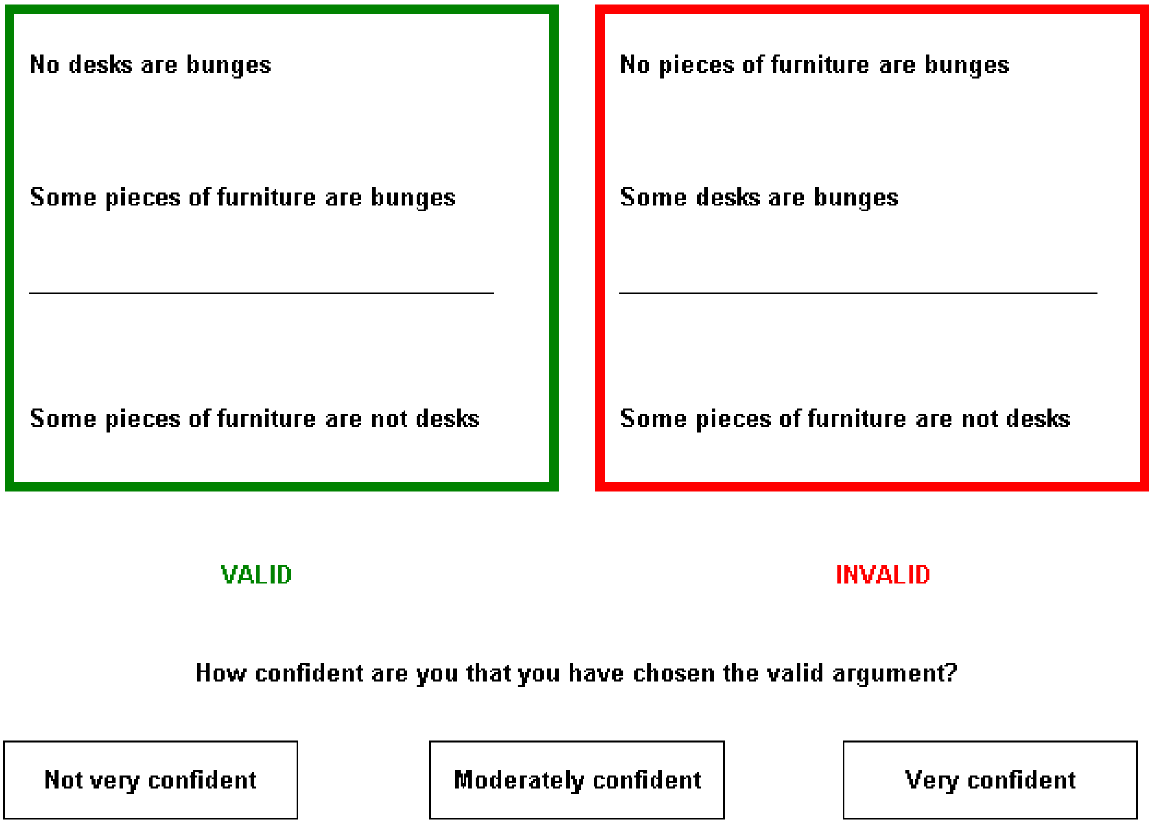


Figure 2.1 Example of the problem setup and response options in Experiment 1.

Measures

On every experimental trial the participant's chosen response (i.e., left argument or right argument) was recorded, as well as whether this response was correct or incorrect. The participant's confidence that he or she chose the valid argument on a scale from 1 (not very confident) – 3 (very confident) was also measured. Response latency (the time between problem presentation and the confidence rating) was also recorded.

Design

The experiment had a one-way design with problem type manipulated within subjects (believable v. unbelievable v. neutral). Every participant received a unique problem set containing identical syllogistic structures which were each assigned problem contents randomly. Within every list all the problems were presented in a randomised order.

2.2.2 Results

Participants were allowed to change their response but only did so in fewer than 2% of the cases. In order to check whether the participants showed a bias for the left or the right argument, the amount of left responses per participant was compared with 24 using a one sample t-test. Participants were not biased in either direction $t(21) = 1.09$, $p = .29$. Two participants were removed from the subsequent analyses because they responded below chance indicating that they did not follow the instructions or engage with the task (including these participants did not change the overall conclusions). Given that accuracy was measured on a proportion scale, homogeneity and normality assumptions were possibly violated. In order to account for this, all proportions were arcsine transformed prior to the analysis (Milligan, 1987). This transformation is used in all further experiments. Reported means and standard errors are the original untransformed values.

Accuracy

To verify that the participants responded significantly above chance, the average proportion of correct responses across participants was compared with arcsine(.50). Overall, participants scored significantly above chance ($M = .69$), $t(19) = 6.48$, $p < .001$.

Motivated reasoning

In order to investigate whether participants engaged in motivated reasoning, a one-way repeated measures ANOVA (problem type: believable v. unbelievable) on proportion correct responses was performed. The analysis revealed that there was no difference in accuracy for believable (M = .69) and unbelievable (M = .71) problems, $F(1, 19) = 0.58$, $p = .46$, $\eta_p^2 = .03$. A neutral item type was also included to investigate whether motivated reasoning in the forced choice reasoning was positive or negative. In order to test for this, a repeated measures ANOVA (problem type: believable v. unbelievable v. neutral) on proportion correct was performed. There was no effect of problem type, $F(2, 38) = 0.71$, $p = .50$, $\eta_p^2 = .04$. Means and standard errors can be found in Table 2.5.

Table 2.5

Experiment 1: Means (Standard Errors) for the Accuracy, Confidence Rating and Latency Analysis

	Believable	**Unbelievable**	**Neutral**
Accuracy	.69 (0.04)	.71 (0.04)	.68 (0.03)
Confidence	2.13 (0.08)	2.17 (0.09)	1.89 (0.07)
Latency	21,977 (1,508)	19,768 (932)	22,198 (1,572)

Note. Accuracy is proportion correct responses, confidence is on a scale from 1 – 3, and latency is measured in ms.

Confidence

Confidence was compared as a function of problem type using one-way repeated measures ANOVA (problem type: believable v. unbelievable v. neutral). The analysis revealed a main effect of problem type, $F(2, 38) = 11.35$, $p < .001$, $\eta_p^2 = .37$. Post-hoc tests revealed that compared to neutral problems, confidence was higher for believable, $t(19) = 3.91$, $p = .001$, and unbelievable, $t(19) = 4.20$, $p < .001$ ones. There was no difference in confidence between believable and unbelievable problems, $t(19) = 0.69$, $p = .50$. Means and standard errors can be found in Table 2.5.

Latency

Response latency (RT) in milliseconds was compared for each problem type using one-way repeated measures ANOVA (problem type: believable v. unbelievable v. neutral). Prior to analysis all RT values were transformed using the natural logarithm. The analysis revealed that RT marginally differed between conditions, $F(2, 38) = 2.90$, $p = .068$, $\eta_p^2 = .13$. Follow-up tests revealed that participants responded significantly more quickly to unbelievable than to believable problems, $t(19) = 2.40$, $p = .027$, and marginally more quickly to unbelievable than to neutral problems, $t(19) = 2.06$, $p = .053$. There was no significant believable–neutral difference, $t < 1$, $p = .90$. Means and standard errors can be found in Table 2.5.

2.2.3 Discussion

A forced choice reasoning method was used to test for motivated reasoning. There was no effect of beliefs on reasoning accuracy. These results are incompatible with motivated reasoning, but compatible with the response bias account of belief bias proposed by Dube et al. (2010). Confidence was higher for believable and unbelievable problems compared to neutral ones, which has been demonstrated in previous work (Shynkaruk & Thompson, 2006). Participants responded significantly quicker to unbelievable problems, which is inconsistent with the claim made by many traditional accounts of belief bias that additional reasoning takes place for syllogisms with unbelievable conclusions (Evans et al., 2001; Thompson, Striemer, Reikoff, Gunter, & Campbell, 2003).

2.3 Experiment 2

The aim of the Experiment 1 was to test for motivated reasoning whilst controlling for response bias. The absence of motivated reasoning is consistent with the response bias account and provides support for Dube et al. (2010). In Experiment 2 our aim was to reintroduce the possibility for response bias in order to find converging evidence for the response bias account (note that traditional accounts also predict response bias alongside motivated reasoning). Given the novelty of the forced choice reasoning method, the question was raised whether this method could result in response bias when it was specifically reintroduced.

Two new problem types were created to allow for a test of response bias. In the new problem types, believability and validity were intentionally confounded, opening up the possibility for participants to base their responses on the believability of the conclusion. The non-conflict problem type consisted of a valid and an invalid syllogism. Conclusion believability was manipulated such that the valid syllogism always had a believable conclusion and that the invalid syllogism always had an unbelievable one. In other words, validity and believability were perfectly positively correlated. For the conflict problems, the valid argument always had an unbelievable conclusion and the invalid argument had a believable conclusion (validity and believability were perfectly negatively correlated; see Table 2.6). The reintroduction of different believability statuses within each trial allowed participants to base their validity decision on their belief in the conclusion. For the non-conflict problems, doing so would result in a high overall proportion of "correct" responses (if for the wrong reason), given that the believable conclusion always followed a valid argument. For the conflict problems, basing the validity judgement on prior beliefs would result in a low overall proportion of correct responses. In other words, if participants judge validity on the basis of believability as all belief bias accounts predict, then proportion "correct" for non-conflict problems should be higher than for conflict problems. As such, the contrast between proportion correct responses for the non-conflict and the conflict problems is a measure of belief bias as response bias. Traditional accounts of belief bias as well as the response bias account predict that participants will show response bias. This prediction is tested in the current experiment.

Table 2.6

Example of the Non-Conflict and Conflict Problem Types

A. Non-conflict	VALID BELIEVABLE: No cigarettes are inexpensive Some addictive things are inexpensive ____________________ Some addictive things are not cigarettes	INVALID UNBELIEVABLE: No cigarettes are inexpensive Some addictive things are inexpensive ____________________ Some cigarettes are not addictive
B. Conflict	VALID UNBELIEVABLE: No addictive things are inexpensive Some cigarettes are inexpensive ____________________ Some cigarettes are not addictive	INVALID BELIEVABLE: No addictive things are inexpensive Some cigarettes are inexpensive ____________________ Some addictive things are not cigarettes

Note. Both syllogisms have conclusions that vary in believability. The left syllogisms is valid, the right one invalid. For the non-conflict problem type (A.) valid syllogisms always have a believable conclusion, invalid syllogisms always have an unbelievable one. For the conflict problem type (B.) valid syllogisms have unbelievable conclusions, invalid syllogisms have believable ones. The presented syllogisms are taken from Evans et al. (1983) and were not used in the present experiment.

2.3.1 Method

Participants

A total of 33 participants took part in exchange for course credit. Six participants were male and 27 were female (age range = 18 – 26, *M* = 20, *SD* = 2).

Materials

The materials were created in exactly the same way as in Experiment 1. The neutral problem type was dropped in favour of the non-conflict and the conflict problem types. The believable and unbelievable problem types used in Experiment 1 were also included in the current study. This resulted in four problem types for a total of 64 trials (16 of each problem type). For believable problems both conclusions were believable, with one argument always being valid and one argument always invalid. For the unbelievable problems both conclusions were unbelievable, one always being valid and one always being invalid. For the non-conflict problems one conclusion was believable and one conclusion was unbelievable, with the added constraint that the believable conclusion was always valid and that the unbelievable conclusion was always invalid. Finally, for the conflict problems one conclusion was always believable and one conclusion was always unbelievable, with the believable conclusion always constrained to be invalid contrary to the unbelievable conclusion, which was always valid. Within each problem type, half of the times the argument on the left was valid and half the times the argument on the right was valid. Item contents were randomized as in Experiment 1. An additional subordinate category member item was added for each of the 16 categories to compensate for the two added problem types. Sixteen extra pseudowords were randomly generated to be used as nonsense middle terms, allowing premise believability to be controlled for.

Procedure, Measures, and Design

The procedure was identical to that of Experiment 1. Accuracy, confidence ratings, and response latency were measured. The experiment employed a one-way within subjects design (problem type: believable v. unbelievable v. non-conflict v. conflict). As in the previous experiment, every participant received an identical list of problems with uniquely randomized item contents. Within every list all the problems were presented in a randomised order.

2.3.2 Results

On each trial, participants were able to change their response from the left syllogism to the one on the right or vice versa, but only did so in 3% of the cases. A one-sample t test comparing proportion of left responses with 32 indicated that the participants were not biased towards the left or the right argument, $t(32) = 1.22$, $p = .23$. Three participants scored below chance for the believable and unbelievable problem types and were removed from all subsequent analyses. Note that we only took performance into account for the believable and unbelievable problems types as they are a true measure of reasoning performance. The conflict and non-conflict problems do not provide a pure measure of reasoning accuracy due to the (intentional) confound between validity and beliefs.

Accuracy

On the whole, the sample reasoned significantly above chance, with the average participant getting 68% correct for the believable and unbelievable problem types, $t(29) = 6.07$, $p < .001$. The current study allows for a test of response bias as well as for a test of motivated reasoning. The relevant contrast for the former test is between the non-conflict and the conflict problem types. For the latter, the contrast of interest is the difference between the believable and the unbelievable problem types.

Response bias

Accuracy was analysed using one-way repeated measures ANOVA (problem type: non-conflict v. conflict). The analysis revealed that there was no significant difference in accuracy for the non-conflict (M = .67) and the conflict (M = .67) problem types, F(1, 29) = 0.02, p = .90, η_p^2 < .01. Means and standard errors can be found in Table 2.7.

Motivated reasoning

Accuracy was analysed using one-way repeated measures ANOVA (problem type: believable v. unbelievable). The analysis revealed that there was no significant difference in accuracy for the believable (M = .67) and the unbelievable (M = .69) problem types, F(1, 29) = 0.042, p = .84, η_p^2 < .01. Means and standard errors can be found in Table 2.7.

Table 2.7

Experiment 2: Means (Standard Errors) for the Accuracy, Confidence Rating and Latency Analysis

	Believable	**Unbelievable**	**Non-conflict**	**Conflict**
Accuracy	.67 (.03)	.69 (.03)	.67 (.04)	.67 (.04)
Confidence	2.00 (0.05)	2.04 (0.06)	1.99 (0.06)	2.03 (0.07)
Latency	16,307 (1,082)	17,074 (931)	16,514 (1,042)	16,183 (913)

Note. Accuracy is proportion correct responses, confidence is on a scale from 1 – 3, and latency is measured in ms.

Confidence

Confidence ratings were analysed using one-way repeated measures ANOVA (problem type: believable v. unbelievable v. non-conflict v. conflict). There was no effect of problem type on confidence, F(3, 87) = 0.47, p = .70, η_p^2 = .02. Means and standard errors can be found in Table 2.7.

Latency

RTs were transformed using the natural logarithm and analysed using one-way repeated measures ANOVA (problem type: believable v. unbelievable v. non-conflict v. conflict).

There was no effect of problem type on latency, $F(3, 87) = 1.37$, $p = .26$, , $\eta_p^2 = .05$. Means and standard errors can be found in Table 2.7.

2.3.3 Discussion

The forced choice reasoning paradigm introduced in Experiment 1 was extended with two additional problem types to allow for a test of response bias and of motivated reasoning. The main finding of Experiment 1 was replicated: beliefs did not affect accuracy, suggesting that participants did not engage in motivated reasoning. This finding is compatible with the response bias account of belief bias (Dube et al., 2010). Incompatible with the response bias account, however, was the finding that belief did not affect performance on the non-conflict and conflict trials. This suggests that participants did not show a response bias even when given the opportunity to do so. This finding requires explanation, because these results imply that belief bias is absent in syllogistic reasoning when studied in a forced choice paradigm.

There are two potential explanations for these results. The first option is that the presentation of the motivated reasoning problem types (believable and unbelievable) alongside the response bias problem types (non-conflict and conflict) caused participants to infer that believability was not a relevant and/or useful problem cue to base a validity judgement on in the current experiment. Given that for half the problems beliefs did not allow participants to make a validity decision (i.e., the believable and unbelievable problem types), they may have chosen to completely ignore beliefs altogether, instead defaulting to content-independent reasoning strategies. This explanation is referred to as the *belief suppression hypothesis*. The second possibility is that presenting syllogisms in a forced choice context triggers a shift in focus from the conclusion to the premises of the argument. In the traditional single-problem presentation paradigm, research suggests that participants typically reason from the conclusion to the premises (Morley, Evans, & Handley, 2004). Participants first evaluate the believability of the conclusion, the assessment of which drives the reasoning process, leading to belief bias. It is possible that the forced choice reasoning paradigm interferes with this process, resulting in an absence of (both components of) belief bias in favour of more structure based reasoning. We call this explanation the *shifted focus hypothesis*. In the remainder of this chapter we test the belief suppression hypothesis (Experiment 3) and the shifted focus hypothesis (Experiment 4).

2.4 Experiment 3

The initial aim of the forced choice reasoning paradigm was to allow for an investigation of response bias and motivated reasoning using a novel SDT-based method. In Experiments 1 and 2, evidence consistent with Dube et al.'s (2010) response bias account was found. Contrary to the clear predictions from all belief bias accounts (including the response bias account) however, response bias did not occur when the opportunity to find it was reintroduced in Experiment 2. This finding suggests that the forced choice method had an impact on the reasoning process. One suggestion was that the method led to the active suppression of prior beliefs when reasoning.

In the current study the belief suppression hypothesis is tested. According to this hypothesis, response bias was absent because of the inclusion of the motivated reasoning test within subjects. Otherwise put, the believable and unbelievable item types revealed to participants that believability is an irrelevant cue for the task at hand. Consequently, participants suppressed the influence of beliefs even for those problems in which it could be used (i.e., the non-conflict and conflict problems). If the belief suppression hypothesis holds, then presentation of only those problems for which beliefs could potentially be considered a useful discriminatory cue (i.e., the non-conflict and conflict problems) should reintroduce response bias. In the current study participants were presented only with non-conflict, conflict and neutral problems. All accounts of belief bias (including the response bias account) predict that accuracy should be higher for the non-conflict compared to the conflict trials due to the application of a belief selection heuristic. As in Experiment 1, the neutral problem type was included to investigate the direction of the belief bias (Evans et al., 2001).

2.4.1 Method

Participants

A total of 23 participants took part in exchange for course credit. Four participants were male and 19 were female (age 18 – 25, M = 20, SD = 2).

Materials

The stimulus list contained 48 problems in total. Sixteen problems were of the non-conflict type, 16 problems were of the conflict type and 16 problems were of the neutral type. All problems were constructed in the same way as in the previous experiments.

Procedure, Measures, and Design

The procedure was completely identical to the procedures of Experiments 1 and 2. The same dependent variables as in Experiments 1 and 2 were measured: accuracy, confidence and latency. A one-way within subjects design was used (problem type: no conflict v. conflict v. neutral). Every participant received an identical list of problems with uniquely randomized item contents. For each participant the problems were presented in an individually randomised order.

2.4.2 Results

On every trial participants were given the opportunity to change their response from one option to the other and back again. In total, participants only switched in 3% of the cases after making their initial selection. Participants were not biased towards the left or the right argument, $t(22) = 0.52$, $p = 0.60$. Two participants responded below chance on the neutral problems and were removed.

Accuracy

Response bias

The remaining participants performed above chance for the neutral problems (M = .73), $t(20) = 5.80$, $p < .001$. To test for belief bias as response bias, accuracy was analysed using a one-way repeated measures ANOVA (problem type: non-conflict v. conflict). The analysis revealed that there was no significant accuracy difference between the non-conflict (M = .71) and conflict (M = .68) problem types, $F(1, 20) = 0.18$, p = .68. The accuracy level of the non-conflict and conflict problem types was compared with a neutral problem type using one-way repeated measures ANOVA (problem type: non-conflict v. conflict v. neutral). There were no significant accuracy differences between the three problem types, $F(2, 40) = 0.32$, p = .73. Means and standard errors can be found in Table 2.8.

Table 2.8

Experiment 3: Means (Standard Errors) for the Accuracy, Confidence Rating and Latency Analysis

	Non-conflict	**Conflict**	**Neutral**
Accuracy	.71 (0.04)	.68 (0.05)	.73 (0.03)
Confidence	2.23 (0.07)	2.27 (0.07)	1.95 (0.07)
Latency	16,449 (1,030)	16,725 (893)	20,009 (1,155)

Note. Accuracy is proportion correct responses, confidence is on a scale from 1 – 3, latency is measured in ms.

Confidence

Confidence ratings were analysed using a one-way repeated measures ANOVA (problem type: non-conflict v. conflict v. neutral). A significant main effect of problem type was found, $F(2, 40) = 18.06$, $p < .001$, $\eta_p^2 = .48$. Pairwise contrasts revealed that confidence was lower for the neutral compared to the non-conflict, $t(20) = 4.97$, $p < .001$, and conflict, $t(20)$

= 4.76, p < .001, problem types. There was no confidence difference between the non-conflict and conflict problem types, t < 1, p = .21. Means and standard errors can be found in Table 2.8.

Latency

Log-transformed RTs were analysed using one-way repeated measures ANOVA (problem type: non-conflict v. conflict v. neutral). The analysis produced a significant main effect of problem type, $F(1, 20) = 8.63$, $p = .001$, $\eta_p^2 = .30$. Pairwise contrasts revealed that participants took longer to respond to the neutral problems compared to the non-conflict, $t(20) = 3.40$, $p = .003$, and the conflict, $t(20) = 3.09$, $p = .006$, problems. Non-conflict and conflict RTs did not differ significantly, t < 1, p = .55. Means and standard errors can be found in Table 2.8.

2.4.3 Discussion

The main aim of Experiment 3 was to test the belief suppression hypothesis. According to this hypothesis, the participants in Experiment 2 did not show response bias because the presence of the motivated reasoning problem types (believable and unbelievable) emphasised that believability was not a relevant cue (i.e., if both conclusions are (un)believable given that one argument is valid and the other is invalid, then it necessarily follows that conclusion believability is unrelated to validity). The absence of belief bias as response bias in the current experiment in the absence of the motivated reasoning problem types suggests that the belief suppression hypothesis is incorrect. Note that in all experiments so far the vast majority of the participants responded significantly above chance, suggesting that the absence of motivated reasoning (E1-E2) and of response bias (E2-E3) and cannot be attributed to floor performance – the participants reason, they just do so in a nontraditional way. Furthermore, the presence of believable and unbelievable contents versus neutral content did have an impact on confidence and latency: participants were more confident and responded quicker to the non-conflict and conflict problems compared to the neutral problems, yet performance did not differ. Taken together these findings suggest that perhaps the forced choice method in its current form encourages a different reasoning style. In the discussion of Experiment 2 we proposed that such an alternative reasoning style may have originated from a shift in focus from the conclusion to the premises. In Experiment 4 the shifted focus hypothesis was tested.

2.5 Experiment 4

According to the shifted focus hypothesis, the mere act of presenting two reasoning problems next to each other fundamentally alters the reasoning process in such a way that participants focus on the premises first, rather than the conclusion. In other words, forced choice reasoning might trigger premise-to-conclusion reasoning, whereas typically, conclusion-to-premise reasoning is the norm (Morley et al., 2004). In addition to this, we suspect that presenting a valid and an invalid argument next to each other reveals the underlying structure. In many cases, for instance, the only difference between the two problems is that the A- and C- terms appear to be swapped around in the premises, with all other things remaining equal. A focus on the structure of the argument rather than the believability of the conclusion might explain why beliefs have no impact and performance is reliably above chance.

In the current experiment we attempted to reintroduce conclusion-to-premise reasoning. If the shifted focus hypothesis is correct, then nudging people towards a conclusion-to-premise reasoning technique should lead to a reintroduction of beliefs in the reasoning process – be it in terms of a response bias or motivated reasoning. We altered the forced choice method by introducing non-simultaneous problem presentation. As in the previous experiments, participants were presented with two syllogisms side by side, however, at the start of each trial, only the conclusions were presented (Figure 2.2).

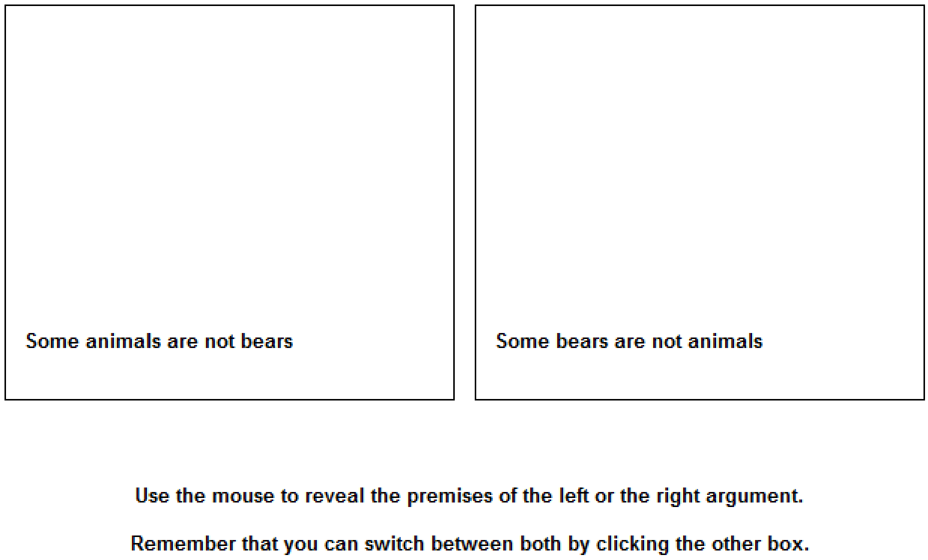


Figure 2.2 Example of the non-simultaneous problem presentation used in Experiment 4, prior to selection.

The participant then had to reason about both syllogisms individually by revealing the premises of each argument by clicking on the corresponding box (Figure 2.3). This ensured that participants processed the conclusions prior to processing the premises. Furthermore, the fact that both syllogisms were never simultaneously visible ensured that a structure based comparison of the premises was impractical. If the non-simultaneous problem presentation reintroduces the effects of belief, we can further investigate whether participants engage in response bias (Dube et al., 2010), motivated reasoning, or both (e.g., Evans et al., 2001).

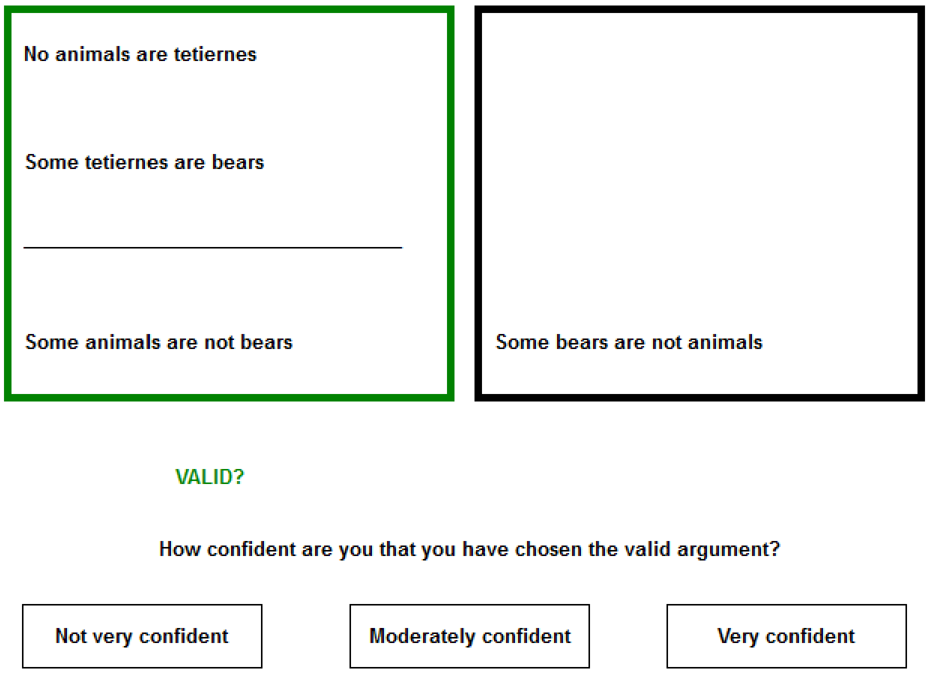


Figure 2.3 Example of the non-simultaneous problem presentation used in Experiment 4, after selection.

2.5.1 Method

Participants

A total of 53 people took place in exchange for course credit (18 male). The mean age was 20 years (*SD* = 3, range = 18 – 31).

Materials

The materials were identical to the ones used in Experiment 2. Every participant received a problem list containing 64 trials with 16 problems of each problem type (i.e., believable, unbelievable, non-conflict, conflict). As in all previous experiments, item contents were randomly assigned for each participant resulting in a unique problem content – problem structure combination, allowing us to control for any unwanted effects of item content.

Procedure, Measures, and Design

The procedure was identical to the procedure in the previous experiments, with the exception of the non-simultaneous problem presentation (see Figures 2.2 and 2.3 for an example). On each trial, participants were presented with two boxes containing only a conclusion. Upon clicking either box, the premises associated with that syllogism appeared, allowing the participant to reason about its validity while the other premises remained hidden. Clicking the non-selected box revealed the premises associated with the other syllogism and made the premises of the first syllogisms disappear (see Figures 2.2 and 2.3). As such, participants could only ever reason about one problem at a time. The participants were allowed to switch between both problems as often as they desired. As in all previous experiments we measured accuracy, confidence ratings on a scale from 1 – 3 and response latency in milliseconds. The experiment employed a one-way within subjects design (problem type: believable, unbelievable, non-conflict, conflict).

2.5.2 Results

Participants were allowed to switch between the left and the right argument as often as they desired. The amount of selection switches was a lot higher than in the previous experiments, with participants swapping between 1 to 2 times per trial on average (M = 1.56, SD = 1.52). This makes sense in light of the fact that participants had to make at least one switch to see both syllogisms. Contrary to the previous experiments, only one participant responded below chance for the believable and unbelievable problems and was removed.

Accuracy

Participants responded significantly above chance for the believable and unbelievable problems (M = .75), $t(51) = 10.34$, $p < .001$. We analysed the accuracy data for the motivated reasoning problems (i.e., the believable and unbelievable problem types) and the response bias problems (i.e., the no conflict and conflict problem types) separately.

Response bias

Accuracy was analysed using a one-way repeated measures ANOVA (problem type: non-conflict v. conflict). The analysis revealed that there was no significant difference in accuracy for the non-conflict (*M* = .77) and the conflict (*M* = .73) problem types, *F*(1, 51) = 0.49, *p* = .49, η_p^2 < .01. Means and standard errors can be found in Table 2.9.

Motivated reasoning

Accuracy was analysed using a one-way repeated measures ANOVA (problem type: believable v. unbelievable). The analysis revealed a main effect of problem type, suggesting that participants achieved higher logical reasoning accuracy for unbelievable (*M* = .77) than for believable (*M* = .72) arguments, F(1, 51) = 4.10, p = .048, η_p^2 = .074. Means and standard errors can be found in Table 2.9.

Table 2.9

Experiment 4: Means (Standard Errors) for the Accuracy, Confidence Rating and Latency Analysis

	Believable	**Unbelievable**	**Non-conflict**	**Conflict**
Accuracy	.72 (.02)	.77 (.03)	.77 (.02)	.73 (.03)
Confidence	2.13 (0.06)	2.24 (0.06)	2.17 (0.06)	2.23 (0.07)
Latency	22,001 (1,147)	21,208 (1,130)	21,080 (1,077)	20,668 (999)

Note. Accuracy is proportion correct responses, confidence is on a scale from 1 – 3, latency is measured in ms.

Confidence

Response bias

Confidence ratings were analysed using a one-way (problem type: non-conflict v. conflict) repeated measures ANOVA. There was no effect of problem type, *p* > .20. Means and standard errors can be found in Table 2.9

Motivated reasoning

Confidence ratings were analysed using a one-way (problem type: believable v. unbelievable) repeated measures ANOVA. Participants were more confident for the unbelievable than for the believable problems, $F(1, 51) = 6.42$, $p = .014$, $\eta_p^2 = .11$. Means and standard errors can be found in Table 2.9.

Latency

Response bias

Log-transformed RTs were analysed using a one way (problem type: non-conflict v. conflict) repeated measures ANOVA. There was no effect of problem type on latency, $p > .81$. Means and standard errors are presented in Table 2.9.

Motivated reasoning

Log-transformed RTs were analysed using a one way (problem type: believable v. unbelievable) repeated measures ANOVA. There was also no effect of problem type on latency, $p > .33$. Means and standard errors are presented in Table 2.9.

2.5.3 Discussion

In Experiment 4 non-simultaneous problem presentation was used to test the shifted focus hypothesis. Compatible with this hypothesis, non-simultaneous problem presentation revealed an effect of belief using the forced choice paradigm. Participants engaged in motivated reasoning, but there was no evidence of response bias (although the means certainly indicated a trend in the expected direction). This finding is incompatible with Dube et al.'s response bias account of belief bias, given that this account predicts exactly the reverse pattern (i.e., the presence of response bias in the absence of motivated reasoning).

With regards to the traditional accounts, the findings also do not fit well with misinterpreted necessity (MN; Evans et al., 1983), metacognitive uncertainty (MU; Quayle

& Ball, 2000) and modified verbal reasoning theory (MVRT; Thompson et al., 2003). All of these accounts do not interpret the logic x belief interaction as originating from motivated reasoning, but rather from a misunderstanding of the logical meaning of necessity (MN), from overloaded working memory resources (MU), or because model search is easier for believable than for unbelievable problems (MVRT). In all of these cases the logic x belief interaction is interpreted as resulting from response bias differing in magnitude for valid and invalid problems, something which cannot possibly drive the accuracy effect given our use of the force choice method explicitly prevented such a response bias from occurring. The findings are more compatible with the following three accounts of belief bias: selective scrutiny (SS; Evans et al., 1982), mental models theory (MMT; Oakhill et al., 1989), and selective processing theory (SPT; Evans et al., 2001; Klauer et al., 2000). Out of the three viable theories, the results are the least compatible with SS given that it would predict both response bias and motivated reasoning, and that only the latter is found. Furthermore, SS does not explicitly suggest that reasoning occurs from the conclusion to the premises; as such there should be no difference in the effects of belief between the simultaneous and the non-simultaneous problem presentation. According to MMT, the key aspect of belief bias in syllogistic reasoning is a motivated search for counterexamples. However, MMT explicitly predicts premise to conclusion reasoning, so it is unclear how the theory can account for the fact that non-simultaneous problem presentation leads to different results than simultaneous problem presentation. SPT, finally, provides a reasonable account of the findings given that it explicitly predicts conclusion to premise reasoning. However, as with SS, SPT also predicts a small response bias which is not found here. The current findings are not fully compatible with any of the extant theories of belief bias.

2.6 General Discussion

The aim of the current chapter was to test both the response bias (Dube et al., 2010; 2011) and the traditional accounts of belief bias using a novel forced choice reasoning method within the SDT framework. In Experiment 4, significant evidence of motivated reasoning was found using a non-simultaneous presentation method. Conversely, in the first three experiments the motivated reasoning (E1-E2) and response bias (E2-E3) components of belief bias were absent. The absence of response bias in Experiment 1 could be taken as evidence compatible with the response bias account, but the subsequent absence of response bias in Experiment 2 suggested that perhaps the simultaneous forced choice reasoning method encourages a different reasoning process compared to more traditional belief bias tasks.

Two potential explanations for the absence of the belief effects were forwarded. According to the belief suppression hypothesis, the presentation of the motivated reasoning problem types alongside the response bias ones caused participants to disregard beliefs completely, given that they were clearly irrelevant in the determination of logical validity. According to the shifted focus hypothesis, presenting two arguments side by side triggered a focus on the structure of the premises rather than the conclusion, leading to an atypical reasoning process different from what is typically expected in a single presentation paradigm. Experiment 3 provided evidence incompatible with the belief suppression hypothesis by showing that simultaneous problem presentation of the response bias problems in the absence of the motivated reasoning problems did not lead to belief bias as response bias. Experiment 4 provided evidence compatible with the shifted-focus hypothesis, suggesting that changing the presentation method to encourage a conclusion-to-premise reasoning strategy resulted in an effect of beliefs on reasoning.

2.6.1 Eliminating Response Bias

The fact that simultaneous problem presentation can be used to eliminate belief bias both in terms of response bias and motivated reasoning whilst still retaining above chance performance is an unexpected and novel finding given the reliability with which the interference of belief with reasoning has been demonstrated countless times since Wilkins (1929). This is particularly interesting considering that many studies have unsuccessfully attempted to eliminate belief bias using various manipulations. For instance, augmented instructions stressing the importance of logical necessity have been shown to reduce, but not eliminate belief bias (Evans, 2000; Evans, Newstead, Allen, & Pollard, 1994; Newstead, Pollard, Evans, & Allen, 1992). The only exception is the case of reasoning with causal conditionals, where response was eliminated using strong instructions, but only for participants of higher cognitive ability (Evans, Handley, Neilens, & Over, 2010). Furthermore, the response bias that was observed was of a smaller magnitude than that in syllogistic reasoning, probably explaining why it was easier to eliminate.

In the traditional syllogistic belief bias paradigm, limiting the amount of response time available has led to the elimination of motivated reasoning in favour of more response bias (Evans & Curtis-Holmes, 2005). However, the validity of all these results is questionable given that these experiments did not employ the appropriate measurement method (i.e., SDT) to quantify response bias and motivated reasoning. It appears that the simultaneous forced choice reasoning method is an interesting new paradigm for the study of deductive reasoning in the absence of both components of belief bias.

2.6.2 Reintroducing Motivated Reasoning

Non-simultaneous problem presentation led to the reintroduction of an influence of beliefs, albeit only in terms of motivated reasoning. This finding is striking for a number of reasons. First, this is–to our knowledge–the first demonstration of motivated reasoning in a forced choice reasoning paradigm, suggesting that the effect is not necessarily a statistical artefact originating from the application of an inappropriate analysis technique (i.e., endorsement rate analysis). Second, the finding of motivated reasoning in the absence of response bias is unexpected and novel, particularly given that the latter merely requires participants to apply a simple belief-heuristic. Even though this finding is potentially compatible with various traditional theories of belief bias, it cannot be fully explained by any one account (see the Discussion of Experiment 4).

The main aim of this chapter was to assess the viability of Dube et al.'s (2010) response bias account using an alternative novel method rooted in the SDT framework. Our results provide evidence incompatible with the idea that belief bias is just a response bias and that motivated reasoning is just an artefact of inappropriate assumptions. These results indicate that a more thorough investigation of both belief bias components is required. We return to the forced choice method in Chapter 4 in which we assessed the impact of individual differences. We now first turn to Chapter 3 in which we employed the ROC method in combination with various traditional experimental manipulations that have been shown to have an impact on belief bias. The primary aim of Chapter 3 is to replicate the absence of motivated reasoning reported by Dube et al. using their method. This is important to ensure that our motivated reasoning evidence which has led us to question the response bias account is not merely due to unforeseen methodological issues inherent to the forced choice paradigm.

Chapter 3: The ROC Method

3.1 Introduction

In Chapter 2 tentative evidence against the response bias account of belief bias was found using a novel forced choice paradigm situated within the SDT framework. This raises the question why our findings clash with those reported by Dube, Rotello, and Heit (2010), who found that response bias can sufficiently account for belief bias. In contrast, our results showed that beliefs resulted in motivated reasoning instead of response bias. This finding was puzzling given that both experiments adopted a methodology situated within the SDT framework. There was a difference, however: Dube et al. gathered confidence ratings to plot receiver operating characteristic (ROC) curves in order to disentangle response bias and reasoning accuracy. We, on the other hand, achieved the same goal by employing a forced choice reasoning approach. Even though the two methods are conceptually equivalent, we cannot exclude the possibility that a difference in the methods has led to the difference in the belief bias effects, as suggested for instance by the presence of motivated reasoning in the absence of response bias in Experiment 4. On a similar note, previous research in the medical imaging domain has shown that the forced choice and ROC procedures can lead to different results (Burgess, 1995).

In the current chapter we report four experiments employing the ROC method. This leaves open two options: a) application of the ROC method will replicate Dube et al.'s findings, suggesting that there is something unique about the forced choice method which leads to motivated reasoning, or b) application of the ROC method will mimic our forced choice findings, suggesting that an alternative difference between Dube et al. and our studies is driving the discrepancy. If the first option turns out to be correct, we need to investigate how and why conceptually equivalent methods within the same framework can lead to clashing results. This finding could potentially put into question the validity of the SDT approach for the study of reasoning. If the second option holds true, a broader focus on the differences between Dube et al.'s experiment and our experiments is necessary.

3.1.1 Belief Bias Manipulations

Various experimental manipulations have been shown to affect belief bias. For instance, decreasing the logical complexity of syllogisms has been shown to eliminate motivated reasoning and to decrease response bias (Newstead, Pollard, Evans, & Allen, 1992). Time taken to respond is has been shown to simultaneously decrease motivated reasoning and increase response bias (Evans & Curtis-Holmes, 2005). Strong or augmented deductive reasoning instructions have been demonstrated to eliminate both response bias and motivated reasoning (Newstead et al., 1992), although later research has shown that rather than eliminate, instructions merely decreased belief bias (Evans, Newstead, Allen, & Pollard, 1994). Item contents can also have unintended effects on syllogistic reasoning and belief bias if the assignment of contents to argument structures is not sufficiently randomised or poorly controlled (Klauer & Singmann, in press). The crucial difference between Dube et al.'s results and the results presented in Chapter 2 could be linked to any of these factors, i.e., a difference in paradigm (forced choice v. ROC), argument complexity, response time taken, instructions, or item contents. We investigate the potential impact of all these factors in four experiments.

3.1.2 The ROC Procedure

To ensure that the use of a different paradigm does not lie at the basis of the inconsistent motivated reasoning findings, we employed the ROC procedure in the next studies. In the traditional conclusion evaluation paradigm, SDT posits that the reasoner compares the strength of the current argument with a criterion. If the argument strength exceeds the criterion, the participant responds "valid", if not, an "invalid" response is given (see Figure 3.1).

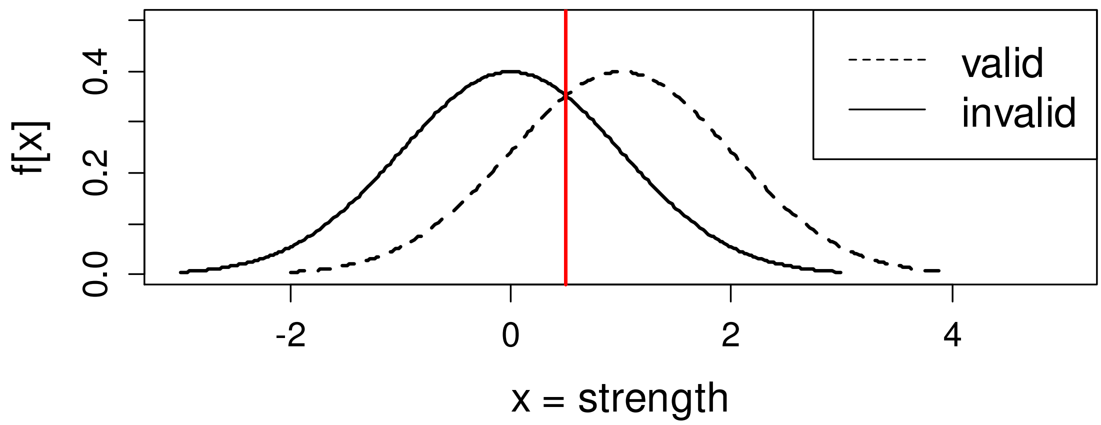


Figure 3.1. Underlying invalid and valid argument strength distributions of the classic binary choice reasoning paradigm according to SDT. "Valid" is responded if argument strength exceeds the response criterion.

One potential issue with the yes/no approach and the single-point sensitivity (d') and bias (c) measures that go with it is that these single point measures require the assumption that the underlying normal distributions of argument strength for valid and invalid arguments have equal variance. Given that this assumption is often violated, it is desirable to enrichen the data by gathering additional observations through manipulation of the response criterion. These additional data points can then be used to verify whether the equal variance assumption holds, and if it does not, additional measures which account for unequal variance can be calculated. In practice this goal is often achieved by gathering confidence ratings (Dube et al., 2010; Macmillan & Creelman, 2005). These extra observations allow for a higher resolution inspection of the underlying distributions. Confidence ratings are used to plot ROC curves, which are plots of hits versus false alarms (i.e., sensitivity) for decreasing levels of the response criterion (i.e., bias). The underlying assumption of using confidence ratings to construct ROCs is that differences in confidence given identical reasoning accuracy reflect internal criterion placement. When reasoning with confidence ratings, the participant must divide the argument strength dimension into

6 bins using 5 response criteria (Figure 3.2), as opposed to two bins using one response criterion in the case of a valid/invalid decision (Figure 3.1). If the argument strength exceeds the most conservative criterion, a high confidence "valid" (i.e., 6) response is given. More liberal responses are linked to lower values on the recoded confidence rating scale. Confidence ratings for "invalid" responses are recoded such that highest confidence "invalid" responses equate to the most liberal response criterion in terms of the underlying argument strength. The reason for this is that responding "invalid" with high confidence implies very low argument strength (i.e., there is only very weak evidence to suggest that the argument is valid), consequently, a response of "1" implies a very liberal response criterion. The relation of the confidence ratings to the argument strength distributions can be seen in Figure 3.2.

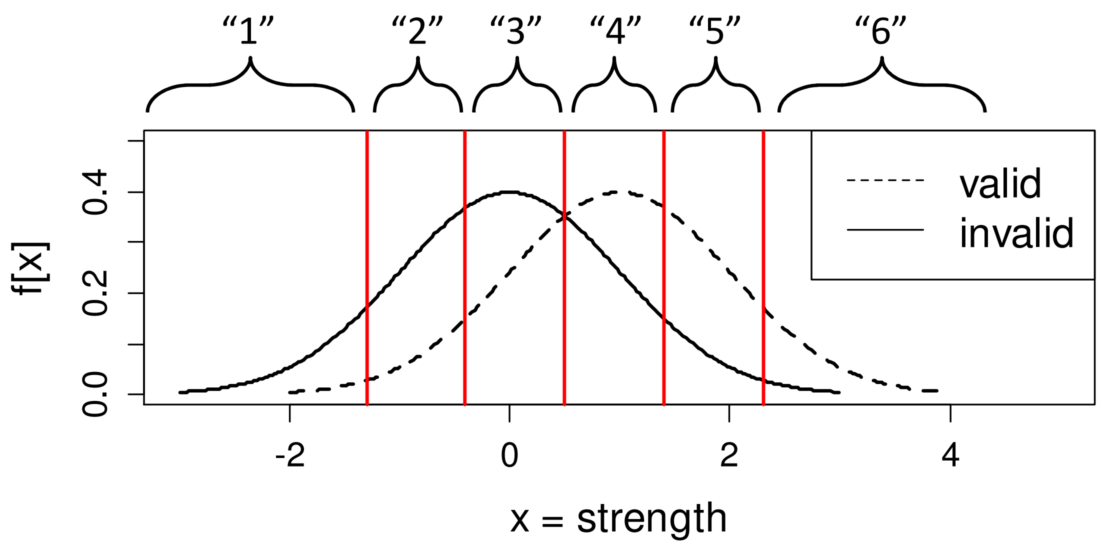


Figure 3.2. Underlying distributions and the 5 criteria of the SDT confidence rating reasoning paradigm. A "6" response equates a high confidence "valid" response. A "1" response equates a high confidence "invalid" response. "5" = medium confidence "valid"; "4" = low confidence "valid" response, "3" = low confidence "invalid" response, "2" = medium confidence "invalid" response.

Figure 3.3 shows how the six confidence ratings can be cumulatively plotted across increasing levels of bias in terms of hits versus false alarms (sensitivity). Due to the cumulative nature of the ROC, and the fact that the density under each distribution equals one, the final point of the ROC necessarily ends up at (1, 1). From a theoretical point of view, optimal performance is achieved when hits are maximized and false alarms are minimized. As such, the closer the resulting curve lies to the upper left corner of the graph, the higher performance. This measure is often formalised by calculating the area under the ROC, A-sub-z (A_z). The explicit relation of the underlying response criteria to the points ultimately plotted in a ROC curve is made evident in Figure 3.3. SDT models can be fit to these ROCs to estimate the mean and variance of the underlying distributions, as well as the location of the 5 criteria.

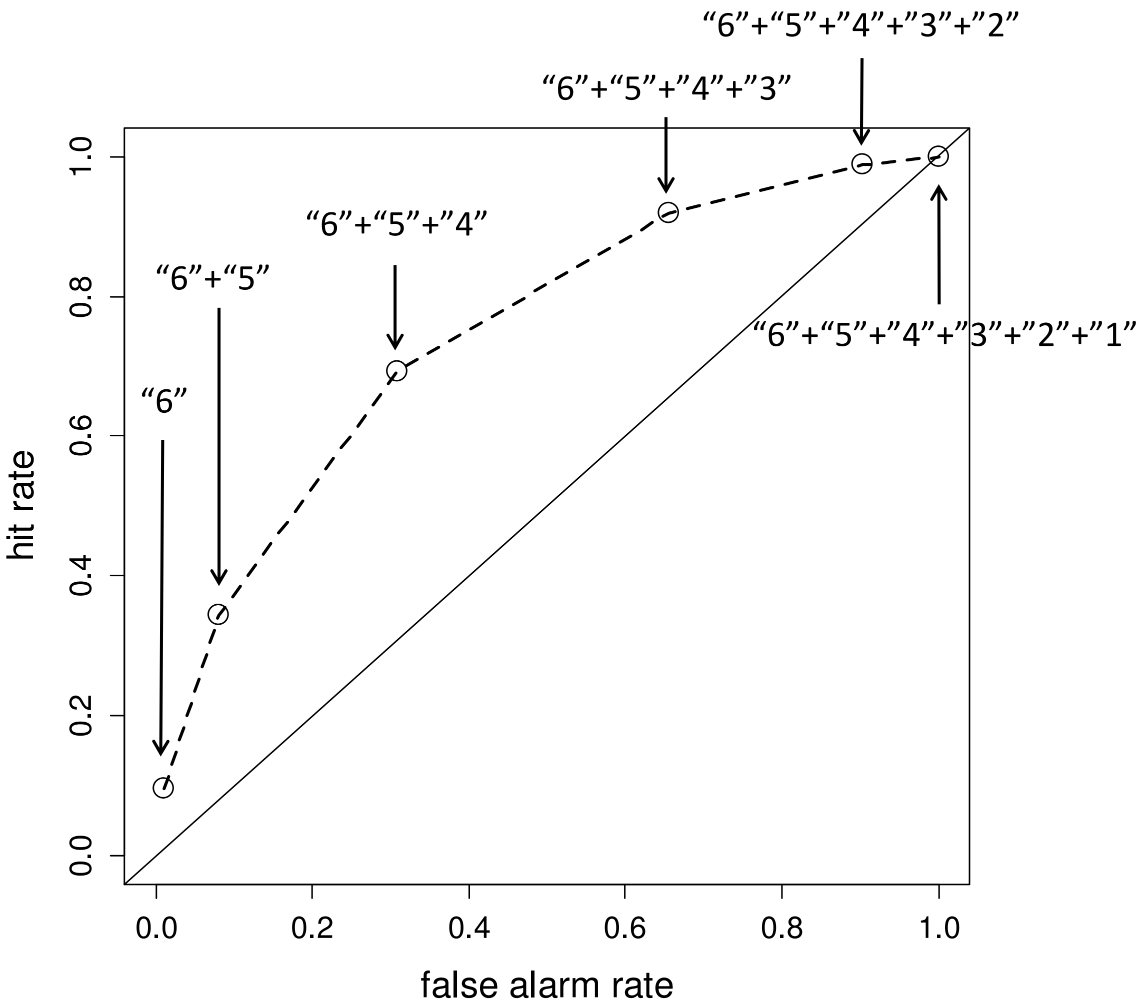


Figure 3.3. Construction of an ROC curve using confidence ratings. Hits are plotted against false alarms for increasing levels of bias.

We now introduce the first experiment in which we attempted to replicate Dube et al.'s finding of response bias in the absence of motivated reasoning using the ROC procedure. We also manipulated syllogism complexity.

3.2 Experiment 5

Dube et al. (2010) argued that belief bias in syllogistic reasoning can be explained in terms of a response bias. Conversely, in Chapter 2 evidence was found that was incompatible with a pure bias account. The main aim of this first experiment was to replicate Dube et al.'s finding using their method. We additionally manipulated argument complexity in order investigate whether their findings generalise to a different class of reasoning problems. Argument complexity was manipulated between subjects by presenting one group of participants with a number of simple (i.e., one-model) syllogisms, whereas a different group was presented with the previously used complex (i.e., multiple-model) syllogisms. We also methodologically improved upon Dube et al.'s (2010) design in multiple ways. Item contents were randomly assigned to syllogistic forms for each participant individually, ensuring that any observed (lack of) differences between conditions are not due to unfortunate random content assignment (e.g., Klauer & Singmann, in press). Finally, given that Dube et al.'s response bias theory is ultimately based on a null-result (i.e., the absence of a difference in accuracy between believability conditions), we increased the available power to detect an effect by doubling the amount of arguments evaluated by each participant.

3.2.1 Theoretical Predictions

If the response bias account holds, there should only be an effect of response bias in both the simple and the complex argument condition. Any observation of motivated reasoning in either condition is clear evidence against the response bias account. The complexity manipulation additionally allows us to distinguish between various traditional accounts of belief bias. Selective scrutiny (SS) predicts more accurate reasoning for unbelievable compared to believable syllogisms because people tend to automatically accept believable conclusions but engage in deductive reasoning for unbelievable ones. This should be true for both simple and complex problems (Evans, Barston, & Pollard, 1983). Misinterpreted necessity (MN) predicts an accuracy effect for complex syllogisms because for conclusions that are indeterminately invalid, people adopt the heuristic of responding solely based on believability. This means that they will tend to be accurate when invalid conclusions are unbelievable but inaccurate when they are believable. This difference in accuracy is not predicted for simple syllogisms because they are determinately invalid (Evans et al., 1983).

Mental Models Theory and Selective Processing Theory both predict an accuracy effect for multiple-model (i.e., complex) syllogisms. According to MMT, unbelievable conclusions cue a motivated search for all possible models, making it likely that an invalid argument will be correctly rejected. Believable conclusions do not cue a search, so that believable invalid arguments are less likely to be rejected (Oakhill, Johnson-Laird, & Garnham, 1989). According to SPT, unbelievable conclusions cue a falsifying strategy so that the single model constructed for an unbelievable invalid argument points to its correct rejection. On the other hand, believable conclusions cue a confirming strategy so that the single model constructed for a believable invalid argument points to its incorrect acceptance (Evans, Handley, & Harper, 2001). The effect on accurate responding described by MMT and SPT does not apply to valid conclusions where there is only a single possible model (i.e., simple syllogisms). The net result is that with complex problems, both MMT and SPT predict an effect of believability on accuracy as a result of differences in the ability to reject invalid conclusions. With simple problems, on the other hand, according to MMT people will always arrive at the same model regardless of differences in reasoning style or strategy. According to SPT, people will arrive at a confirmatory model for valid syllogisms with believable conclusions (leading to correct acceptance) but fail to arrive at a disconfirmatory model for valid syllogisms with unbelievable conclusions (also leading to correct acceptance). Thus, according to both theories, believability is predicted to have no effect on accuracy for the simple problems.

3.2.2 Method

Participants

A total of 91 undergraduate psychology students from Plymouth University participated in exchange for partial fulfilment of a course requirement. Fifteen participants were male and 76 were female (age 18 – 43, M = 20, SD = 4).

Design

Logical validity (valid v. invalid argument) and conclusion believability (believable v. unbelievable) were manipulated within subjects, and argument complexity (simple v. complex syllogisms) was manipulated between subjects.

Materials and Measures

For each participant, we created a unique list of 64 syllogisms (32 valid and 32 invalid) by randomly assigning 64 item contents to the available syllogistic structures. This ensured that item content had the potential to be different in every validity x believability cell for each participant, allowing us to control for content effects. Item contents were identical to the ones used in the experiments presented in Chapter 2. Every list contained 32 valid syllogisms and 32 invalid syllogisms. The simple syllogisms were of the one-model type using all four figures and all quantifiers bar the "Some…not" one (see Appendix A for an overview). These were sampled from Evans, Handley, Harper, & Johnson-Laird (1999). The complex syllogisms were of the multiple-model type and taken from Dube et al. (2010). Believability was manipulated by making the conclusions definitionally true or false. This was achieved by combining object categories with category members. In the simple argument condition, conclusions used the "No"- or "Some"-quantifiers. Believable "No"-conclusions consisted of a category with a member from a different category (e.g., no tools are trout). Unbelievable ones consisted of a category with one of its members (e.g., no tools are hammers). Believable "Some"-conclusions consisted of a category with one of its members (e.g., some tools are hammers). Unbelievable ones consisted of a category with one of its non-members (e.g., some tools are trout). In the complex argument condition conclusions used the "Some…not"-quantifier. Believable conclusions featured the category followed by one of its members (e.g., some tools are not hammers). Unbelievable conclusions featured a category member followed by its category (e.g., some hammers are not tools). We used nonsensical middle terms to control for premise believability. Each list was generated by combining the structures, item contents and middle terms according to the constraints outlined above. Examples of all problem types can be found in Table 3.1. We measured endorsement rates (participant responded "valid" or "invalid"), confidence ratings (1 = not confident at all – 3 = very confident), and response latency in milliseconds.

Table 3.1

Experiment 5: Examples of Simple and Complex Reasoning Problems

		Valid	Invalid
Simple	**Believable**	All bananas are queels	Some cars are remoxtions
		No queels are reptiles	All remoxtions are ducks
		No bananas are reptiles	No cars are ducks
	Unbelievable	Some spaniels are mips	All Fiats are punes
		All mips are fruits	No punes are insects
		Some fruits are spaniels	Some insects are Fiats
Complex	**Believable**	No saws are veemers	Some haddocks are curges
		Some tools are veemers	No curges are fish
		Some tools are not saws	Some fish are not haddocks
	Unbelievable	No birds are pinds	No spears are blans
		Some pinds are parrots	Some weapons are blans
		Some parrots are not birds	Some spears are not weapons

Note. In the actual data set figure and mood were not confounded with problem type. P1 = first premise, P2 = second premise, C = conclusion.

Procedure

Participants were tested on individual computers in small groups. Upon entering the room they were randomly assigned to the simple (n = 47) or complex (n = 44) condition. The instructions were taken from Dube et al. (2010; exp. 2) and presented on the screen. Participants were instructed to assume the premises were true, to judge whether the conclusion necessarily followed, and to rate their confidence. After completing four practice trials, the participants solved the remaining 64 syllogisms. Syllogisms were presented one at a time with response options (s = Valid, k = Invalid) shown at the bottom the screen. After each validity judgment, participants indicated their confidence on a scale from 1 (not confident) to 3 (very confident).

3.2.3 Results

Data treatment

We analysed the data both in terms of endorsement rates and by fitting SDT models to ROCs. We derived SDT indexes of accuracy and bias to disentangle motivated reasoning from response bias (Dube et al., 2010). Validity judgements and confidence ratings were combined into six response bins, from 6 (high confidence valid response) to 1 (high confidence invalid response). ROC curves were constructed from the transformed confidence ratings for each individual and condition. Accuracy (A_z) and bias (c_a) parameters were derived from these ratings using Systat 12 (Macmillan & Creelman, 2005). A_z is a measure of the area under the ROC curve. Given that the ROCs are an estimation of the underlying cumulative probability density function, the range of Az is 0 – 1, with .5 indicating chance performance. The parameter c_a is a measure of the average response criterion adjusted for unequal variance for the valid and invalid argument strength distributions. More negative values indicate a more liberal response criterion (i.e., an increased likelihood of responding "valid"). Proportion measures (endorsement rates and A_z) were arcsine transformed to conform with the assumptions of ANOVA. Response latencies were transformed using the natural logarithm and consequently analysed using ANOVA.

Endorsement

In order to verify whether the participants showed the traditional belief bias effect, endorsement rates were submitted to a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (complexity: simple v. complex) mixed ANOVA, with the first two factors manipulated within subjects and the last one between subjects. The analysis revealed a main effect of validity showing that participants endorsed valid arguments more than invalid arguments, $F(1, 89) = 398.66$, $p < .001$, $\eta_p^2 = .82$. This effect interacted with complexity suggesting that the validity effect was larger in the simple arguments condition, $F(1, 89) = 97.80$, $p < .001$, $\eta_p^2 = .51$. Participants also endorsed believable arguments more than unbelievable arguments, $F(1, 89) = 10.09$, $p = .002$, $\eta_p^2 = .10$. A main effect of complexity indicated that participants in the complex condition accepted more arguments than participants in the simple condition, $F(1, 89) = 12.70$, $p = .001$, $\eta_p^2 = .13$. Logical validity and conclusion believability interacted in the traditional way, suggesting that the validity effect was larger for unbelievable than for believable arguments, $F(1, 89) = 5.47$, $p = .022$, , $\eta_p^2 = .06$. Logic, belief and complexity also interacted, $F(1, 89) = 9.82$, $p = .002$, , $\eta_p^2 = .10$. Follow-up tests revealed that the logic x belief interaction was significant in the complex condition, $F(1, 46) = 10.22$, $p = .003$, $\eta_p^2 = .19$, but not in the simple condition, $F(1, 46) < 1$, $p = .47$. Means and standard errors can be found in Table 3.2.

Table 3.2

Experiment 5: Endorsement Rates per Complexity Condition.

	Valid		**Invalid**	
Complexity	**Believable**	**Unbelievable**	**Believable**	**Unbelievable**
Simple	.91 (.02)	.87 (.03)	.11 (.03)	.06 (.02)
Complex	.78 (.03)	.75 (.04)	.60 (.04)	.41 (.03)

Note. Means (Standard Errors).

SDT

ROCs (Figure 3.4) were constructed to derive accuracy and bias for each participant. Accuracy (A_z) was analysed to test for motivated reasoning. Bias (c_a) was analysed to test for response bias. Both variables were analysed separately.

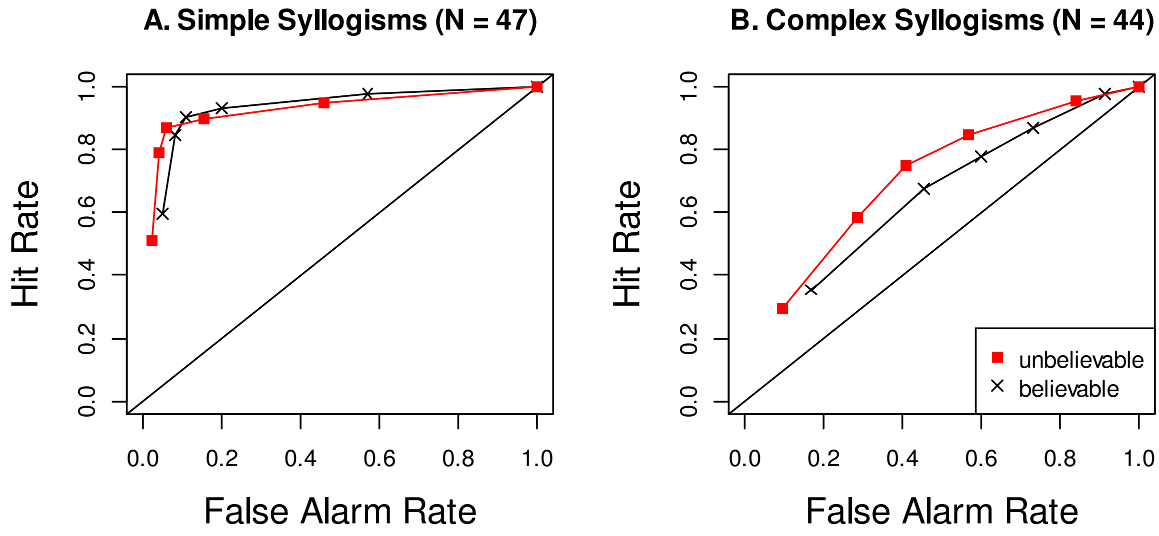


Figure 3.4 Experiment 5: A) ROC for the believable and unbelievable conditions in the simple condition. B) ROC for the believable and unbelievable conditions in the complex condition.

Motivated reasoning

Accuracy (A_z) was analysed using a 2 (believable v. unbelievable) x 2 (simple v. complex) mixed ANOVA, with the first factor manipulated within subjects and the second factor manipulated between subjects. The analysis revealed a main effect of beliefs on accuracy, suggesting that reasoning accuracy was higher for unbelievable than for believable arguments, $F(1, 89) = 6.90$, $p = .010$, $\eta_p^2 = .07$. The analysis also revealed a main effect of complexity showing that logical accuracy was higher in the simple than in the complex condition, $F(1, 89) = 69.68$, $p < .001$, $\eta_p^2 = .52$. Believability and complexity interacted, $F(1, 89) = 5.51$, $p = .021$, $\eta_p^2 = .058$. Follow up tests revealed that participants in the complex condition showed motivated reasoning, $t(43) = 2.65$, $p = .011$, whereas those in the simple condition did not, $t(46) = 0.34$, $p = .73$. Means and standard errors can be found in Table 3.3.

Response bias

Response bias (c_a) was analysed using a 2 (believable v. unbelievable) x 2 (simple v. complex) mixed ANOVA, with the first factor manipulated within subjects. The analysis revealed a main effect of believability, suggesting that participants responded “valid” more to believable arguments than to unbelievable arguments, $F(1, 89) = 10.33$, $p = .002$, $\eta_p^2 = .10$. There was also a main effect of complexity on bias, suggesting that participants were more liberal in the complex condition, $F(1, 89) = 65.31$, $p < .001$, $\eta_p^2 = .42$. The two factors did not interact, $F(1, 89) = 0.39$, $p = .54$, $\eta_p^2 < .01$. Means and standard errors can be found in Table 3.3.

Table 3.3

Experiment 5: SDT parameters per Complexity Condition

	Simple problems		**Complex problems**	
Measure	**Believable**	**Unbelievable**	**Believable**	**Unbelievable**
Accuracy (A_z)	.89 (.02)	.90 (.02)	.64 (.02)	.70 (.02)
Bias (c_a)	-0.01 (0.06)	0.18 (0.07)	-0.53 (0.07)	-0.25 (0.07)

Note. Means (Standard Errors).

Latency

We analysed latency by submitting log-transformed response times to a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (argument complexity: simple v. complex) mixed ANOVA with repeated measures on the first two factors. Validity and complexity interacted, $F(1, 89) = 25.48$, $p < .001$, $\eta_p^2 = .22$. Follow up tests tests indicated that for simple arguments participants took longer to respond to valid (M = 10,595 ms) than to invalid (M = 10,008 ms) arguments, $t(46) = 3.10$, $p = .003$. For complex arguments the reverse held as participants took longer to respond to invalid (M = 12,417 ms) than to valid (M = 11,299 ms) arguments, $t(43) = 3.86$, $p < .001$. No other effects were significant, all ps > .16. Means and standard errors can be found in Table 3.4.

Table 3.4

Experiment 5: Response Times per Complexity Condition.

	Valid		Invalid	
Complexity	**Believable**	**Unbelievable**	**Believable**	**Unbelievable**
Simple	10,703 (592)	10,488 (503)	10,045 (597)	9,972 (567)
Complex	11,538 (612)	11,061 (520)	12,338 (617)	12,496 (586)

Note. Means (Standard Errors) are expressed in milliseconds.

3.2.4 Discussion

Both the endorsement rate analysis and the SDT analysis converged in showing that participants engaged in motivated reasoning. This is incompatible with Dube et al.'s (2010) finding that belief bias is just a response bias and that beliefs do not impact upon accuracy when the appropriate measurement method (SDT) is used. With simple syllogisms, the believability effect was consistent with a pure response bias. With complex syllogisms, however, unbelievable conclusions produced more accurate judgments.

The two observations of motivated reasoning using two different paradigms within the SDT framework (i.e., Experiment 4 and the current experiment) raise the question why Dube et al.'s findings differ from ours. At first glance, the response time data do not seem to support a motivated reasoning account given that conclusion believability had no impact on the response times. Similar findings have been taken as evidence against motivated reasoning (Thompson, Newstead, & Morley, 2011; Thompson, Striemer, Reikoff, Gunter, & Campbell, 2003), however, these apparently discrepant findings can be explained in two ways. First, response time does not necessarily equate to reasoning time. Thompson et al. argue that participants respond before a self-imposed deadline elapses. It is plausible that the overall time taken to respond is comparable across believability conditions, but that the amount of time allocated to reasoning and rationalising differs in such a way that participants are more engaged in reasoning in the unbelievable condition and in rationalising for the believable condition. Second, research of response times broken down by reasoning competence has suggested that the effects typically observed in syllogistic reasoning research are an artefact of aggregating across multiple subgroups of reasoners each applying different reasoning strategies (Stupple, Ball, Evans, & Kamal-Smith, 2011). We will return to this point later on.

Research has shown that limiting the amount of response time available affects belief bias by eliminating motivated reasoning in favour increased response bias (Evans & Curtis-Holmes, 2005). Given Thompson et al.'s (2003) argument that reasoners respond within an arbitrarily set response deadline, it possibly follows that the observed difference between our experiment and Dube et al.'s experiment is due to random sampling differences in the participants' tendency to set a strict response deadline. If, for some reason, Dube et al.'s participants chose to respond more quickly than ours, this might explain the absence of motivated reasoning in their study. We test this hypothesis in Experiment 6.

3.3 Experiment 6

The aim of the current experiment is to investigate whether a response time deadline leads to the disappearance of motivated reasoning in favour of response bias within the SDT framework. Evans and Curtis-Holmes (2005) have demonstrated that participants put under time pressure engage in more response bias and no motivated reasoning. In the current experiment participants were randomly assigned to an untimed or a speeded response time condition. If the hypothesis holds that the discrepancies between Dube et al. and our study stem from a difference in response time taken, we should find motivated reasoning in the untimed condition but no motivated reasoning in the limited time condition. From the traditional accounts we would derive the predictions that limiting response time should lead to the elimination of motivated reasoning, an increase in response bias, and an overall decrease in accuracy compared to the untimed condition. From the response bias only account we predict that limiting response time should lead to a decrease in overall accuracy and an increase in response bias, because limiting the possibility to engage in effortful reasoning should shift the participants towards a less effortful strategy such as a simple belief-acceptance heuristic. The response bias account predicts that motivated reasoning is absent in all conditions.

3.3.1 Method

Participants

A total of 86 Plymouth University undergraduate psychology students participated in exchange for course credit (34 male, age: range 18 – 49, M = 20, SD = 4).

Design, Materials and Measures

Logical validity (valid v. invalid argument) and conclusion believability (believable v. unbelievable conclusion) were manipulated within subjects, and time limit (10 s response time limit v. no response time limit) was manipulated between subjects. The materials were constructed in an identical manner as in the complex condition of Experiment 5. We measured endorsement rates, confidence ratings (1 – 3), response latency, and amount of missing trials in the speeded condition.

Procedure

The procedure was identical to the one used in Experiment 5, with two exceptions: a) all participants were presented with complex syllogisms and b) response time was manipulated between subjects by randomly assigning half the participants to a speeded condition. Under speeded instructions participants had to respond within a 10 second response time deadline. This deadline was chosen on the basis of previous research (Evans & Curtis-Holmes, 2005). On each trial, a red bar at the top of the screen ran out as the available response time decreased. If no response had been made after ten seconds the participant would be urged to respond more quickly next time and the trial would be discarded. All 64 trials were presented in a completely random order for each participant. The unspeeded condition was identical to the complex condition in Experiment 5.

3.3.2 Results

Data treatment

Endorsement rates, SDT accuracy and bias parameters, and latencies were analysed as in Experiment 5. Less than 2% of the responses were made after the deadline. These trials were eliminated from all further analyses.

Endorsement

Endorsement rates were submitted to a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (time limit: no limit v. speeded) mixed ANOVA, with repeated measures on the first two factors. The analysis revealed a main effect of validity suggesting that participants accepted valid arguments more than invalid arguments, $F(1, 84) = 67.3$, $p < .001$, $\eta_p^2 = .45$. Participants also responded "valid" to believable arguments more than to unbelievable arguments, $F(1, 84) = 30.0$, $p < .001$, $\eta_p^2 = .26$. Validity and believability interacted, $F(1, 84) = 4.39$, $p = .039$, $\eta_p^2 = .05$. The interaction was opposed to the expected pattern such that the validity effect appeared to be larger in the believable condition than in the unbelievable condition. No other effects were significant, all *p*s > .12. Means and standard errors can be found in Table 3.5.

Table 3.5

Experiment 6: Endorsement Rates per Time Limit Condition.

	Valid		Invalid	
Time Limit	**Believable**	**Unbelievable**	**Believable**	**Unbelievable**
Speeded	.74 (.03)	.54 (.04)	.57 (.03)	.40 (.03)
No Limit	.78 (.03)	.61 (.04)	.53 (.03)	.46 (.03)

Note. Means (Standard Errors).

SDT

ROCs (Figure 3.5) were constructed to derive accuracy (A_z) and bias (c_a) for each participant. We tested for motivated reasoning and response bias separately.

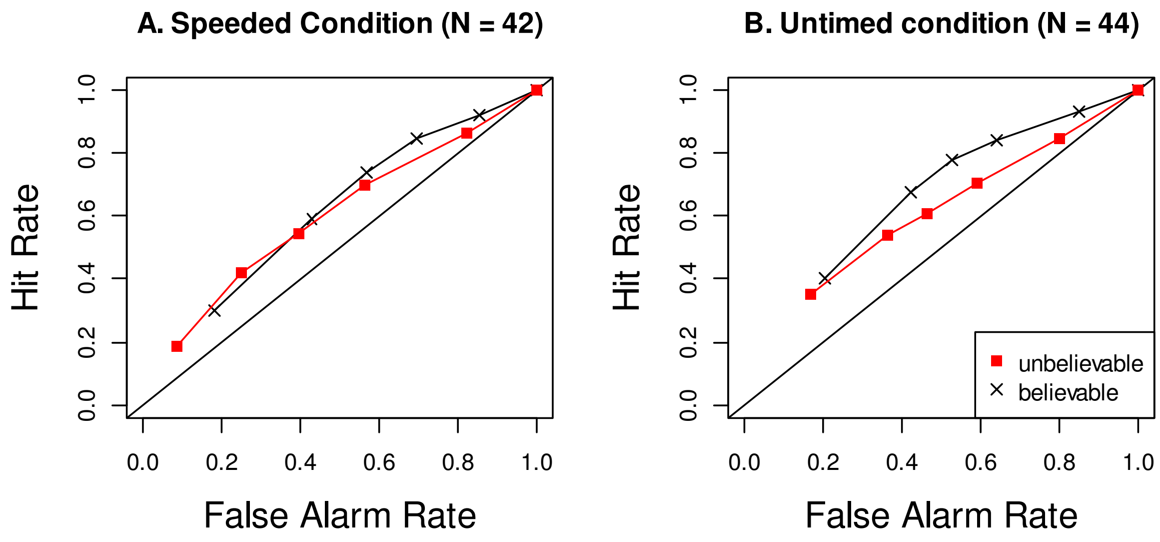


Figure 3.5. A) ROC for the believable and unbelievable conditions in the speeded condition of Experiment 6. B) ROC for the believable and unbelievable conditions in the untimed condition of Experiment 6.

Motivated reasoning

Accuracy was analysed using a 2 (believable v. unbelievable) x 2 (untimed v. speeded) mixed ANOVA, with repeated measures on the first factor. The analysis did not produce any significant effects. Accuracy was not affected by beliefs, $F(1, 84) = 1.99$, $p = .16$, time limit, $F(1, 84) = 1.31$, $p = .26$, or the interaction between both, $F(1, 84) < 1$, $p = .49$. Means and standard errors can be found in Table 3.6

Response bias

Bias was analysed using a 2 (believable v. unbelievable) x 2 (untimed v. speeded) mixed ANOVA, with repeated measures on the first factor. The analysis revealed a main effect of beliefs on bias in such a way that participants were more likely to endorse believable than unbelievable arguments, $F(1, 84) = 32.8$, $p < .001$, $\eta_p^2 = .28$. No other effects were significant, all *ps* > 14. Means and standard errors can be found in Table 3.6

Table 3.6

Experiment 6: SDT parameters per Time Limit Condition

	Speeded condition		**Untimed condition**	
Measure	**Believable**	**Unbelievable**	**Believable**	**Unbelievable**
Accuracy (A_z)	.61 (.02)	.60 (.03)	.66 (.02)	.59 (.03)
Bias (c_a)	-0.44 (0.07)	0.09 (0.07)	-0.42 (0.07)	-0.11 (0.07)

Note. Means (Standard Errors).

Latency

Response times were analysed using a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (time limit: no limit v. speeded) mixed ANOVA with repeated measures on the first two factors. Participants responded more quickly to valid than to invalid arguments, $F(1, 84) = 28.0$, $p < .001$, $\eta_p^2 = .25$. Participants in the speeded condition responded more quickly than participants in the untimed condition, $F(1, 84) = 43.0$, $p < .001$, $\eta_p^2 = .34$. The two factors also interacted, suggesting that validity effect was larger in the untimed condition than in the speeded condition, $F(1, 84) = 9.1$, p

= .003, η_p^2 = .10. No other effects were significant, all *p*s > .21. Means and standard errors can be found in Table 3.7.

Table 3.7.

Experiment 6: Response Times per Time Limit Condition.

	Valid		Invalid	
Time Limit	**Believable**	**Unbelievable**	**Believable**	**Unbelievable**
Speeded	6,578 (390)	6,697 (378)	6,831 (578)	6,930 (552)
No Limit	10,227 (381)	10,078 (369)	11,660 (565)	11,960 (540)

Note. Means (Standard Errors) are expressed in milliseconds.

3.3.3 Discussion

The aim of the current experiment was to investigate whether differences in response time taken potentially underlie the discrepancies between the data reported by Dube et al. (2010) and our experiments. The results of the current experiment appear puzzlingly inconsistent with earlier data. The finding of a response bias in the absence of motivated reasoning appears to be compatible with the response bias account and incompatible with traditional accounts of belief bias. Limiting the amount of response time drove the participants towards quicker responding but had no impact on overall reasoning quality or response bias. This is surprising based on previous research showing that limiting response time has an impact on belief bias (Evans & Curtis-Holmes, 2005; although this could be explained by the fact that this study did not use the more appropriate SDT method). The fact that response time had no impact on response bias or overall reasoning accuracy is odd from the response bias account point of view. According to this account participants engage in a nondescript reasoning process in combination with a belief-acceptance heuristic. Limiting response time should at least have a negative impact on accuracy and possibly increase the overall belief bias observed.

The current experiment leaves us with data inconsistent with all belief bias accounts, although it is more compatible with the response bias account than with the traditional accounts. So far our experiments resulted in two significant observations of motivated reasoning (Experiments 4 and 5) and one absence of motivated reasoning (the current experiment). Motivated reasoning was also absent in two additional experiments in Chapter 1, but this was probably due to the use of the simultaneous forced choice reasoning method. In the next experiment we further attempt to resolve these apparent inconsistencies by investigating the effect of instructions on belief bias within the ROC paradigm.

3.4 Experiment 7

Instructional manipulations have been shown to impact response bias and motivated reasoning. Evans, Handley, Neilens, and Over (2010) investigated the effect of pragmatic (i.e., weak) instructions on belief bias in causal conditionals. Unlike the traditional instructions given in deductive reasoning tasks, weak instructions encourage participants to respond on the basis of whether the conclusion follows from the premises or not, without explicitly instructing the participants to assume the truth of the premises. Results showed that pragmatic instructions led to increased response bias compared to deductive instructions (see also George, 1995; Stevenson & Over, 1995).

The opposite approach was taken by Newstead et al. (1992), who compared the effect of augmented (or strong) instructions on syllogistic belief bias. These strong instructions emphasized the necessity requirement of logical validity. Their findings showed that motivated reasoning was eliminated under strong instructions. Later studies nuanced this finding by confirming that strong instructions reduced the impact of beliefs on reasoning (Evans, Newstead, Allen, & Pollard, 1994). Taken together, these findings are reminiscent of the forced choice data presented in Chapter 2: our forced choice results indicated that conditions which emphasised the logical structure of the argument (i.e., the simultaneous presentation method used in Experiments 1 – 3) eliminated any effects of beliefs – both in terms of response bias and motivated reasoning. In contrast, a condition in which a structural focus was explicitly avoided (i.e., the non-simultaneous presentation method used in E4) led to a belief effect – specifically in terms in motivated reasoning.

In the current experiment we investigated the effect of pragmatic instructions on belief bias in syllogistic reasoning. One straightforward prediction is that the weak instructions should lead to increased response bias compared to a standard instructions condition (Evans et al., 2010). This also raises the question how beliefs might impact on motivated reasoning. One possibility is that the weak instructions will create a condition more conducive to motivated reasoning in parallel to the non-simultaneous problem presentation method.

We randomly assigned participants to a standard or weak instructions condition. The standard instructions condition is identical to the one used in all previously reported experiments, including Dube et al.'s (2010). This condition allows us to potentially replicate the findings of Experiments 4 and 5 (if motivated reasoning is found) or Experiment 6 and Dube et al.'s study (if no motivated reasoning is found). In the weak instructions condition, participants were merely told to respond whether a conclusion followed from the premises or not. No explicit mentions of logic, validity, or necessity were made.

3.4.1 Method

Participants

A total of 66 Plymouth university undergraduate psychology students and volunteers drawn from the paid participant pool were recruited in exchange for course credit or a small fee (24 male, age: range = 18 – 64, M = 28, SD = 11).

Design, Materials, and Measures

Logical validity (valid v. invalid argument) and conclusion believability (believable v. unbelievable conclusion) were manipulated within subjects, and instruction type (weak v. standard) was manipulated between subjects. The materials were constructed in an identical manner as in Experiments 5 and 6. We measured endorsement rates, confidence ratings (1 – 3), and response latency

Procedure

The procedure for the standard instructions condition was identical to the complex condition in Experiment 5 and the untimed condition in Experiment 6. In the weak instructions condition, participants were presented with the following instructions:

In this experiment, we are interested in people's reasoning. For each question, you will be given some information above a line. If you judge that the information below the line follows from this, you should answer "Follows". If you judge that the information below the line doesn't follow, you should answer "Doesn't follow". Please respond as quickly as possible and simply follow your gut feeling. It is important that you follow your intuition.

In the standard instructions condition on each reasoning trial the response options "valid" and "invalid" were available. In contrast, in the weak instructions condition the response options "follows" and "doesn't follow" were presented to allow participants to engage in a less formal reasoning style. Confidence ratings (1 – 3) were collected after each response.

3.4.2 Results

Data treatment

Endorsement rates, SDT (accuracy and bias) parameters, and response latencies were analysed as in Experiments 5 and 6.

Endorsement

Endorsement rates were submitted to a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (instruction type: weak v. standard) mixed ANOVA with repeated measures on the first two factors, instruction type being a between-subjects factor. Participants accepted valid arguments more than invalid arguments, $F(1, 64) = 47.2$, $p < .001$, $\eta_p^2 = .43$. Participants also responded "valid" more to believable arguments than to unbelievable arguments, $F(1, 64) = 24.6$, $p < .001$, $\eta_p^2 = .28$. No other effects approached significance, all *p*s > .18. Means and standard errors can be found in Table 3.8.

Table 3.8.

Experiment 7: Endorsement Rates per Instructions Condition.

	Valid		Invalid	
Instructions	**Believable**	**Unbelievable**	**Believable**	**Unbelievable**
Weak	.76 (.03)	.57 (.05)	.56 (.05)	.40 (.04)
Standard	.83 (.03)	.63 (.05)	.58 (.05)	.35 (.04)

Note. Means (Standard Errors).

SDT

We constructed ROCs (Figure 3.6) to which we fit the SDT model for each participant to derive accuracy (A_z) and bias (c_a). We tested for motivated reasoning and response bias separately.

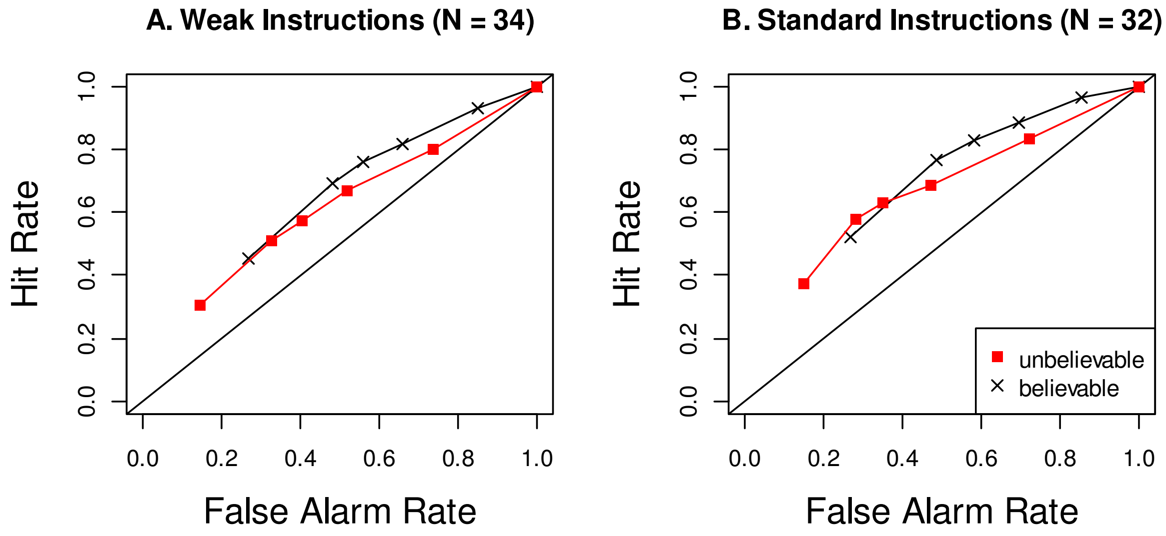


Figure 3.6. A) ROC for the believable and unbelievable conditions in the weak instructions condition of Experiment 7. B) ROC for the believable and unbelievable conditions in the standard instructions condition of Experiment 7.

Motivated reasoning

Accuracy was analysed using a 2 (believable v. unbelievable) x 2 (weak instructions v. standard instructions) mixed ANOVA with repeated measures on the first factor. There were no effects of belief or instructions on accuracy, all *ps* > .24. Means and standard errors can be found in Table 3.9.

Response bias

Response bias was analysed using a 2 (believable v. unbelievable) x 2 (weak instructions v. standard instructions) mixed ANOVA with repeated measures on the first factor. The analysis revealed a significant main effect of beliefs on the response criterion, suggesting that a more liberal response criterion was adapted for believable than for unbelievable arguments, $F(1, 64) = 25.5$, $p < .001$, $\eta_p^2 = .29$. No other effects approached significance, all *ps* > .41. Means and standard errors can be found in Table 3.9.

Table 3.9

Experiment 7: SDT parameters per Instructions Condition

	Weak instructions		**Standard instructions**	
Measure	**Believable**	**Unbelievable**	**Believable**	**Unbelievable**
Accuracy (A_z)	.64 (.03)	.61 (.03)	.68 (.03)	.65 (.03)
Bias (c_a)	-0.45 (0.10)	0.03 (0.10)	-0.61 (0.10)	0.01 (0.11)

Note. Means (Standard Errors).

Latency

We analysed the log transformed response times using a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (instructions: weak v. standard) mixed ANOVA with repeated measures on the first two factors. The analysis only resulted in a significant main effect of validity suggesting that participants responded more quickly to valid than to invalid arguments, $F(1, 64) = 25.2$, $p < .001$, $\eta_p^2 = .28$. No other effects were significant, all *p*s > .13. Means and standard errors can be found in Table 3.10.

Table 3.10

Experiment 7: Response Times per Instructions Condition.

	Valid		**Invalid**	
Instructions	**Believable**	**Unbelievable**	**Believable**	**Unbelievable**
Weak	10,125 (1,057)	10,311 (991)	11,321 (1,281)	10,893 (1,165)
Standard	12,250 (1,057)	11,310 (1,021)	13,918 (1,032)	13,096 (1,201)

Note. Means (Standard Errors) are expressed in milliseconds.

3.4.3 Discussion

We attempted to replicate the motivated reasoning effect demonstrated in Experiments 4 and 5 after we were unsuccessful in doing so in Experiment 6. Participants in the current experiment showed response bias, but no motivated reasoning. We created a weak instructions condition which we hypothesised would result in increased response bias and potentially result in more motivated reasoning. Surprisingly, there was no effect of instructions whatsoever. This runs counter to previous findings which have demonstrated that manipulating instructions affects belief bias (Evans et al., 1994; 2010; Newstead et al., 1992). The fact that presenting participants with pragmatic instructions stressing people to follow their gut feeling led to response patterns which were virtually identical is consistent with the probability heuristics model (PHM; Chater & Oaksford, 1999). According to PHM, participants are not engaged in deduction, but rather attempt to determine whether the conclusion is less uncertain than the premises using a set of five fast and frugal heuristics. This can explain why participants in the pragmatic instructions condition demonstrated similar levels of logical competence as those in a standard instructions condition, even though they were stressed to follow their intuition or gut feeling: perhaps this is what participants are doing all along.

Taken together, the results of Experiments 4 – 7 are confusing. Using the non-simultaneous forced choice method we found one case of motivated reasoning (Experiment 4). Using the ROC method we also found one case of motivated reasoning (Experiment 5), but two cases in which motivated reasoning appeared to be fully absent (Experiments 6 and 7). A potential (if unsatisfying) explanation for the inconsistent findings is that differences in participant motivation levels between experiments are the crucial difference. It seems plausible that a lack of motivation will lead to lower levels of motivated reasoning, whereas such an effect will not have an impact on response bias because it requires less effort due to its heuristic nature. It could be argued that a potential proxy of motivation is the amount of time participants invest in a task, with less motivated participants taking less time to reason per trial. We explored the relationship between motivated reasoning and average response time using a scatterplot. In all further analyses we collapsed across instructions, given the lack of instructional effects on any of the dependent variables. We calculated a motivated reasoning index (MRI) by subtracting accuracy in the believable condition from accuracy in the unbelievable condition for each participant (higher values indicated more motivated reasoning) and plotted it against average response time taken per trial (Figure 3.7).

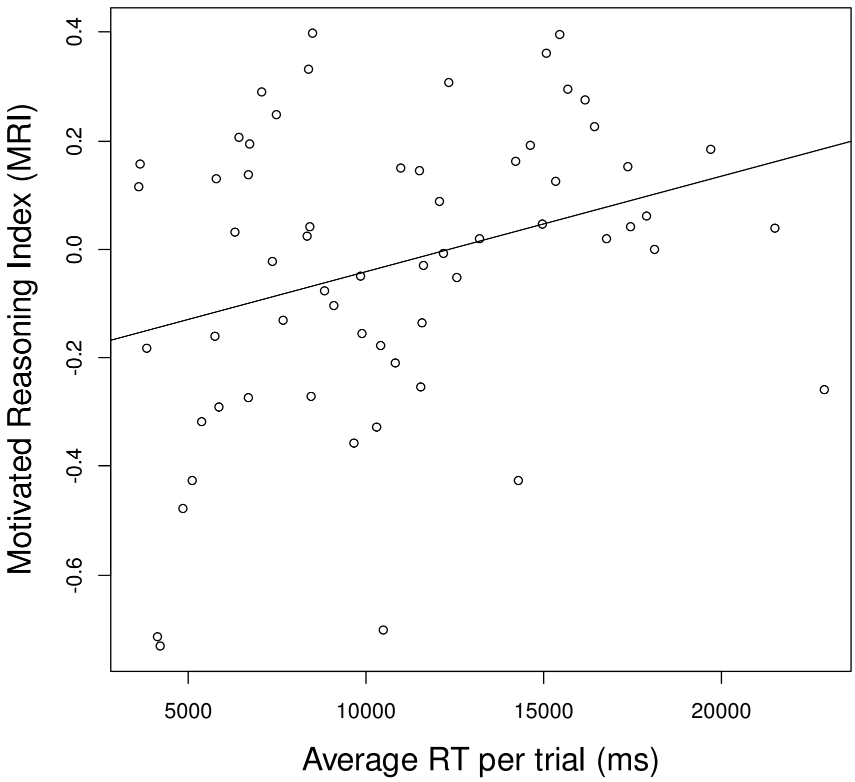


Figure 3.7. Scatterplot of the average latency per trial v. motivated reasoning.

Figure 3.7 demonstrates that longer latencies were significantly correlated to higher degrees of motivated reasoning, $r(61) = .32$, $p = .011$. Visual inspection of the scatterplot in combination with the moderate correlation between average latency and motivated reasoning suggests that more motivated participants (i.e., participants who responded more slowly on average) often showed positive levels of motivated reasoning, with less motivated participants (i.e., quicker responders) showing none, or even reversed motivated reasoning (albeit with much greater variance). To further explore this hypothesis we investigated the degree of motivated reasoning for the top 25% slowest responders only ($M_{\text{latency}} > 15{,}028$ ms). A one sample t test comparing the MRI with 0 suggested that this group engaged in motivated reasoning (M_{MRI} = 0.10), $t(16) = 2.40$, $p = .029$. Interestingly, for the 25% quickest responders ($M_{latency}$ < 6,818 ms) a trend towards reversed motivated reasoning occurred (M_{MRI} = -0.15), $t(16) = -1.93$, $p = .07$. In line with our hypothesis that slower responders were more motivated, a two sample t test of accuracy between the 25% slowest and 25% quickest responders revealed that the slowest reasoners performed significantly better overall ($A_z = .74$) than the quickest reasoners ($A_z = .55$), $t(32) = 5.03$, $p < .001$.

These findings suggest that motivation as approximated by average response time taken was moderately correlated to the degree of motivated reasoning. This is important because it indicates that perhaps the discrepancies in motivated reasoning between our own Experiments (and equally so, between Dube et al.'s study and ours) could be due to individual differences. It is plausible that in Experiment 5 we sampled a larger group of motivated versus unmotivated participants, whereas in Experiments 6 and 7 there were fewer highly motivated reasoners. Interestingly, the negative motivated reasoning exhibited by the unmotivated reasoners has to our knowledge never been reported in the belief bias literature. It is an important finding however, because it suggests that ignoring individual differences by averaging across motivated and unmotivated participants can create the illusion that motivated reasoning is absent (the negative and positive motivated reasoning effects cancel each other out). These data suggest that individual differences should not be ignored and that they provide an interesting avenue for further research. A similar suggestion has been made by Stupple et al. (2011), who analysed latency patterns in syllogistic belief bias in function of reasoning aptitude. Their findings showed that the aggregate RT pattern commonly observed is in fact an artefact caused by averaging across qualitatively different subgroups of reasoners. In Experiment 8 we address the individual differences hypothesis more formally by investigating cognitive ability.

3.5 Experiment 8

The aim of the current experiment was to investigate the impact of individual differences in cognitive ability on belief bias. Cognitive ability is a correlate of working memory capacity (WMC; Conway, Kane, & Engle, 2003) which is associated with reasoning ability (Capon, Handley, & Dennis, 2003). Research on causal conditional reasoning has shown that high cognitive ability participants are able to resist belief bias when instructed to do so compared to a group of lower ability participants (Evans et al., 2010). In syllogistic belief bias it has been shown that the logic x belief interaction (traditionally considered a proxy for motivated reasoning) was present only in a group of high ability participants (Newstead, Handley, Harley, Wright, & Farrely, 2004). This suggests that cognitive ability is a potential key determinant of motivated reasoning. We tested this hypothesis in the current study by comparing a higher and a lower cognitive ability group. If our hypothesis holds that cognitive ability is a determinant of motivated reasoning, we predict that motivated reasoning occurs in the higher ability group only.

We also manipulated response time available as in Experiment 6. If response time available has an impact on belief bias as demonstrated by Evans and Curtis-Holmes (2005), then a drop in motivated reasoning should be observed in the higher ability – limited time group, but response bias should increase regardless of cognitive ability. In contrast, according to the response bias account, motivated reasoning should be absent in all conditions. From the traditional accounts, MMT predicts that working memory capacity may be a prerequisite for conducting the motivated search for counterexamples resulting in better reasoning accuracy for unbelievable compared to believable problems (Oakhill et al., 1989). In contrast, the metacognitive uncertainty account (Quayle & Ball, 2000) makes the opposite prediction that WMC is linked to an absence of the logic x belief interaction, although not through motivated reasoning but rather due to the application of a belief-heuristic for indeterminately invalid problems used by lower-ability participants only. The other traditional accounts do not predict a link between cognitive ability and motivated reasoning.

3.5.1 Method

Participants

Eighty-five undergraduate psychology students from Plymouth University participated for course credit (15 male, age: range 17 – 52, M = 22, SD = 6).

Design, Materials, and Measures

Logical validity (valid v. invalid argument) and conclusion believability (believable v. unbelievable) were manipulated within subjects. Time limit condition (speeded v. unlimited) and cognitive ability (high v. low) were between-subjects factors. The problems were complex syllogisms constructed in the same way as in the previous experiments. Cognitive ability was measured using part 1 of the AH4 Group Test of General Intelligence which contains 65 verbal or numerical questions (Heim, 1970). The test was administered in small groups of five or less. The test consists of ten self-paced practice items followed by 65 test problems of which the participant has to complete as many as possible in the span of ten minutes. Newstead et al. (2004) have shown that scores on part 1 of the AH4 are related to logical performance on a variety of deductive reasoning tasks. As in all previous experiments we measured endorsement rates, confidence ratings on a scale from 1 – 3, and response latency in milliseconds.

Procedure

The procedure was identical to that of Experiment 6, with the exception that participants completed the cognitive ability test before completing the reasoning test.

3.5.2 Results

Data treatment

Endorsement rates, SDT accuracy and bias parameters, and latencies were analysed as in Experiments 5 – 7. Less than 2% of the responses were made outside of the deadline. These were removed prior to the analysis. Participants were assigned to a higher- (n = 41) or lower ability (n = 44) group on the basis of an above- or below median score on the cognitive ability test (*Mdn* = 38).

Endorsement

Endorsement rates were analysed using a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (time limit: speeded v. unlimited) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first two factors. Time limit and cognitive ability were manipulated between subjects. The analysis revealed a main effect of validity suggesting that participants accepted valid arguments more than invalid arguments, $F(1, 81) = 66.3$, $p < .001$, $\eta_p^2 = .45$. Participants also accepted believable arguments more than unbelievable arguments, $F(1, 81) = 16.7$, $p < .001$, $\eta_p^2 = .17$. A main effect of time limit demonstrated that participants in the untimed condition responded "valid" more often than participants in the speeded condition, $F(1, 81) = 4.7$, $p = .034$, $\eta_p^2 = .05$. Logical status and time limit interacted in such a way that the validity effect was larger in the untimed condition, $F(1, 81) = 15.4$, $p < .001$, $\eta_p^2 = .16$. Logical status and cognitive ability marginally interacted, suggesting that the validity effect was larger for the higher ability group, $F(1, 81) = 3.5$, $p = .066$, $\eta_p^2 = .04$. Time limit and cognitive ability also marginally interacted suggesting that the higher ability – untimed group responded "valid" more often than the higher ability – speeded group, whereas there was no such difference between the lower ability groups, $F(1, 81) = 3.1$, $p = .084$, $\eta_p^2 = .04$. There was a three-way interaction between logic, belief, and time limit, $F(1, 81) = 14.8$, $p < .001$, $\eta_p^2 = .16$. The interaction revealed that in the untimed condition, the logic x belief interaction operated in the traditional manner (i.e., with a larger validity effect for unbelievable arguments), whereas for the timed condition the direction of the interaction was reversed. No other effects were significant, all *p*s > .15. Means and standard errors can be found in Table 3.11.

Table 3.11

Experiment 8: Endorsement Rates per Time Limit Condition and Cognitive Ability Group

	Valid		**Invalid**	
Group	**Believable**	**Unbelievable**	**Believable**	**Unbelievable**
Higher cognitive ability				
Speeded	.73 (.04)	.46 (.07)	.48 (.05)	.42 (.06)
Untimed	.83 (.04)	.73 (.06)	.57 (.05)	.32 (.05)
Lower cognitive ability				
Speeded	.70 (.04)	.49 (.07)	.57 (.05)	.51 (.06)
Untimed	.70 (.04)	.67 (.06)	.50 (.05)	.39 (.05)

Note. Means (Standard Errors).

SDT

We constructed ROCs (Figure 3.8) to which we fit the SDT model for each participant to derive accuracy (A_z) and bias (c_a). We tested for motivated reasoning and response bias separately.

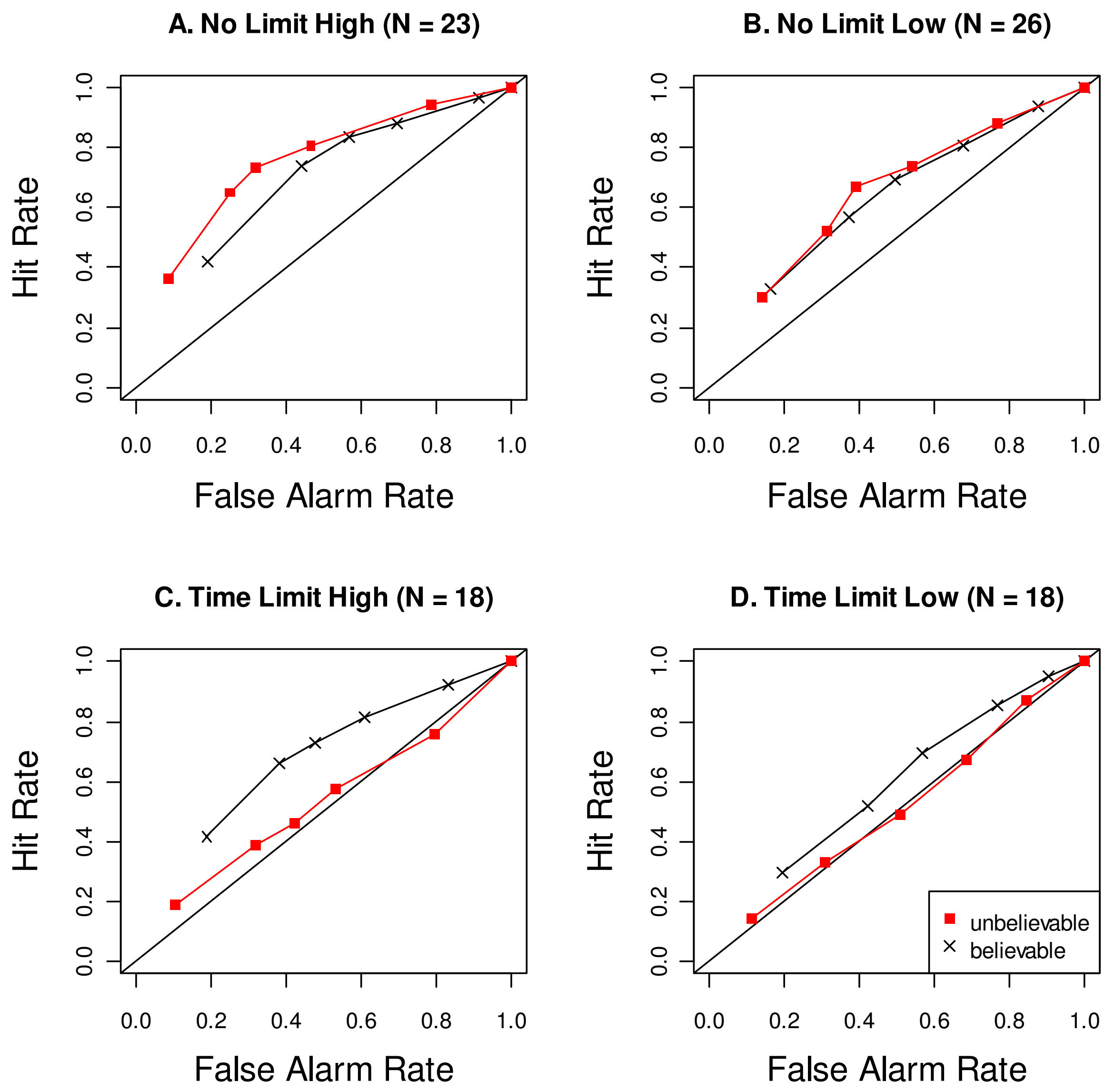


Figure 3.8. Experiment 8: Top row: ROCs for the no time limit condition for the higher (A) and lower (B) cognitive ability groups. Bottom row: ROCs for the time limit condition for the high (C) and low (D) cognitive ability groups.

Motivated reasoning

Accuracy was analysed using a 2 (believable v. unbelievable) x 2 (speeded v. unspeeded condition) x 2 (higher cognitive ability v. lower cognitive ability) mixed ANOVA, with repeated measures on the first factor. Accuracy was higher for the higher ability than the lower ability group, $F(1, 81) = 4.8$, $p = .031$, $\eta_p^2 = .06$. A main effect of time limit suggested that accuracy was higher in the untimed than in the speeded condition, $F(1, 81) = 8.5$, $p = .005$, $\eta_p^2 = .10$. The analysis revealed an interaction between time limit and believability suggesting that in the untimed condition participants appeared to reason better for unbelievable than for believable arguments, whereas the opposite held true in the speeded condition, $F(1, 81) = 8.4$, $p = .005$, $\eta_p^2 = .09$. Follow-up tests revealed that the apparent difference was marginally significant in both the speeded, $t(35) = 2.01$, $p = .052$, and the untimed condition, $t(48) = 1.86$ $p = .069$. There was also a marginal three-way interaction between believability, time limit, and cognitive ability, $F(1, 81) = 3.2$, $p = .076$, $\eta_p^2 = .04$. A separate believability x time limit analysis for the higher and the lower ability groups revealed that there was a significant belief x time limit interaction for the higher ability group, $F(1, 39) = 12.4$, $p = .001$, $\eta_p^2 = .24$, but not for the lower ability group, $F < 1$, $p = .46$. Follow-up tests on the belief x time limit interaction in the higher ability group revealed that motivated reasoning occurred in the untimed condition, $t(22) = 2.46$, $p = .022$, and that reversed motivated reasoning in the timed condition, $t(17) = 2.44$, $p = .026$. Means and standard errors can be found in Table 3.12.

Response bias

Response bias was analysed using a 2 (believable v. unbelievable) x 2 (speeded v. unspeeded condition) x 2 (higher cognitive ability v. lower cognitive ability) mixed ANOVA, with repeated measures on the first factor. The analysis revealed a significant main effect of belief showing that participants were more liberal in the believable than in the unbelievable condition, $F(1, 81) = 19.0$, $p < .001$, $\eta_p^2 = .19$. There was also a marginal interaction between time limit and cognitive ability suggesting criterion placement was more liberal in the untimed – higher ability group than in the speeded – higher ability group, whereas no such difference appeared to be present in the lower ability groups, $F(1, 81) = 2.9$, $p = .094$, $\eta_p^2 = .034$. Follow-up tests revealed that for the higher ability group the unspeeded – speeded difference was significant, $t(39) = 2.4$, $p = .022$, whereas this was not the case for the lower ability group, $t(42) = 0.1$, $p = .95$. Means and standard errors can be found in Table 3.12.

Table 3.12

Experiment 8: SDT parameters by Time Limit Condition and Cognitive Ability Group

Group	Speeded		Untimed	
	Believable	**Unbelievable**	**Believable**	**Unbelievable**
Higher cognitive ability				
Accuracy (A_z)	.68 (.03)	.54 (.06)	.66 (.03)	.72 (.04)
Bias (c_a)	-0.28 (0.12)	0.17 (0.09)	-0.63 (0.13)	-0.10 (0.16)
Lower cognitive ability				
Accuracy (A_z)	.56 (.03)	.51 (.05)	.63 (.03)	.63 (.04)
Bias (c_a)	-0.31 (0.08)	-0.01 (0.12)	-0.22 (0.11)	-0.07 (0.13)

Note. Means (Standard Errors).

Latency

We submitted the log-transformed response times to a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (time limit: speeded v. untimed) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first two factors. The analysis revealed a main effect of logical status suggesting that participants responded more quickly to valid than to invalid arguments, $F(1, 81) = 16.3$, $p < .001$, $\eta_p^2 = .17$. There was also a main effect of time limit showing that participants responded more quickly in the speeded than in the untimed condition, $F(1, 81) = 47.6$, $p < .001$, $\eta_p^2 = .37$. Logic interacted with time limit suggesting that the validity effect was larger in the untimed condition, $F(1, 81) = 8.4$, $p = .005$, $\eta_p^2 = .09$. Follow up tests revealed that participants responded significantly more quickly to valid than to invalid arguments in the untimed condition, $t(48) = 4.07$, $p < .001$. The difference was marginally significant in the speeded condition, $t(35) = 1.94$, $p = .06$. No other effects reached significance, all *ps* > .22. Means and standard errors can be found in Table 3.13.

Table 3.13

Experiment 8: Response Times by Time Limit Condition and Cognitive Ability Group

	Valid		Invalid	
Group	**Believable**	**Unbelievable**	**Believable**	**Unbelievable**
Higher cognitive ability				
Speeded	6,434 (838)	6,554 (990)	6,576 (1,173)	6,792 (1,076)
Untimed	10,982 (742)	11,171 (876)	12,208 (1,037)	12,829 (951)
Lower cognitive capacity				
Speeded	7,140 (838)	6,993 (990)	7,168 (1,173)	7,051 (1,076)
Untimed	12,384 (742)	12,376 (876)	14,351 (1,037)	14,101 (951)

Note. Means (Standard Errors). Response times are expressed in milliseconds.

3.5.3 Discussion

The aim of the current experiment was to investigate the effect of individual differences in cognitive ability on belief bias. Motivated reasoning was engaged in only by the higher cognitive ability participants. This finding is incompatible with the response bias account of belief bias as introduced by Dube et al. (2010). The lower ability group did not show motivated reasoning. As shown previously by Evans and Curtis-Holmes (2005), limiting the available response time eliminated motivated reasoning in this higher ability sample. The finding that neither cognitive ability nor response time limit had an impact on the degree of belief bias as response bias is incompatible with the results reported by Evans and Curtis-Holmes. The finding of reversed motivated reasoning in the speeded high ability condition might appear counterintuitive in light of most of the research reported in the literature, but it echoes the reverse motivated reasoning pattern demonstrated by the quick responders in Experiment 7 suggesting that it might be an effect which has previously been overseen due to aggregation. These findings seem to suggest that ignoring individual differences might be a key driver behind the inconsistencies reported in many of the previous studies of belief bias, including the ones reported in this dissertation so far.

3.6 General Discussion

In the current chapter we aimed to further our understanding of the response bias and motivated reasoning components of belief bias using the ROC method. More specifically, we tried to address any key factors that may have differed between Dube et al.'s (2010) experiment and ours: materials, argument complexity, time taken to respond, instructions, and cognitive ability. In Experiment 5, clear evidence compatible with motivated reasoning was found, suggesting that at least under some conditions, belief bias is not just response bias – regardless of the employed SDT method (i.e., ROC v. forced choice). Experiment 5 also furthered our understanding of belief bias by showing that motivated reasoning occurred only for complex, but not simple syllogisms – as predicted by traditional belief bias accounts such as mental models theory (Oakhill et al., 1989) and selective processing theory (Evans et al., 2001; Klauer et al., 2000), but not by selective scrutiny (Evans et al., 1983).

Response bias occurred under both simple and complex argument conditions, although it was smaller in the former case. In Experiment 6 we did not replicate the motivated reasoning observed in the first experiment. A speeded-task manipulation had no effect on response bias or motivated reasoning. These findings were compatible with the response bias account and puzzling in light of our earlier motivated reasoning findings. In Experiment 7 we investigated the effect of pragmatic instructions on response bias and motivated reasoning. As in Experiment 6, we failed to establish motivated reasoning in the basic instructions condition. Even more puzzling was the finding that pragmatic instructions did not have an impact on response bias or motivated reasoning, suggesting that belief bias as response bias might be even more pervasive than previously assumed.

An exploratory analysis of motivated reasoning as a function of response time taken revealed that those who took longer to respond were more prone to motivated reasoning, with quicker responders engaging in no – or even reversed motivated reasoning. In Experiment 8, finally, we built on this finding by formally investigating the effect of cognitive ability on belief bias. Our findings suggested that motivated reasoning occurred only for a group of higher cognitive ability participants. Contrary to Experiment 6, the speeded-task manipulation eliminated motivated reasoning in this higher ability group.

A host of additional manipulations on belief bias were also investigated. Previous research on complexity has led to inconsistent results, with the logic x belief interaction for single model problems occasionally appearing (Glinsky & Judd, 1994) and disappearing (Newstead et al., 1992; Klauer et al., 2000). Our findings suggested that beliefs do not affect reasoning accuracy for simple syllogisms. Limiting the amount of response time had no discernible impact in Experiment 6, in contrast to previous studies in which it led to increased response bias (Evans, Handley, & Bacon, 2009) and decreased motivated reasoning (Evans & Curtis-Holmes, 2005). In fact, the finding that there was no effect of limiting response time seems to suggest that perhaps the untimed group was reasoning at a subpar level. In Experiment 8 we controlled for poor performance by measuring cognitive ability. This time, the speeded manipulation led to decreased motivated reasoning, although it had no impact on the magnitude of response bias. Interestingly, for the high ability subgroup the time pressure resulted in a reversed motivated reasoning effect, a novel finding in the belief bias literature as far as we know. A similar finding appeared in the quick responding subgroup found in the exploratory analysis of Experiment 7. One potential explanation of this finding is that the high ability group attempted to engage in motivated reasoning using a difficult effortful (Deutsch, Kordts-Freudinger, Gawronski, & Strack, 2009) falsification strategy as proposed by MMT and SPT, but that they failed to complete this process when confronted with a time limit – be it experimentally manipulated (E8) or self-imposed (E7). In contrast, the less effortful confirmatory reasoning strategy presumably engaged for the believable problems did not suffer from such a time limit, leading to preserved performance and the “reversed motivated reasoning” effect seen here. Note that aggregating across standard and reversed motivated reasoners might give the impression that there is no effect of beliefs on accuracy whatsoever, whereas in fact, the vast majority of participants’ reasoning performance is affected by accuracy.

We also collected latency data across all experiments. Previous research has demonstrated three response time effects: participants typically respond more quickly to valid than to invalid arguments and more quickly to unbelievable than to believable arguments. The two factors also interact in such a way that for invalid problems participants respond more quickly to unbelievable than to believable problems, whereas such a difference is absent for valid problems (Thompson et al., 2003). We replicated the validity effect on response latency in Experiments 5 – 8, but there was no effect of beliefs on response time. One potential explanation for the reliable validity effect is that valid arguments are more fluently processed, leading to reduced response times (Morsanyi & Handley, 2012). The absence of an effect of beliefs on response time might seem puzzling in light of earlier findings. However, Stupple et al. (2011) investigated belief bias latency data as a function of reasoning aptitude by comparing the RT patterns for a group of poor, good, and great reasoners. Their data showed that the typically observed effects reported in aggregate analysis are probably an artefact resulting from inappropriate aggregation across qualitatively different subgroups. These findings suggest a) that perhaps latency is not the optimal dependent variable of interest for tackling the belief bias debate and b) that ignoring individual differences and aggregating across groups might lead to inconsistent findings when analysing other, more important dependent variables such as reasoning accuracy, response bias, and motivated reasoning.

Based on the findings reported so far, we propose an individual differences account of belief bias. According to this account, cognitive ability is a prerequisite for motivated reasoning, but not response bias. Although this may be a preliminary conclusion to draw at this point, this account explains the differences in motivated reasoning observed between Experiments 4 through 8, suggesting that random differences in the proportion of higher versus lower cognitive ability participants determined the results. One potential explanation is that in Experiments 4 and 5 we randomly sampled a larger group of high ability participants compared to Experiments 6 and 7. We now turn to Chapter 4 in which we investigated the individual differences account of belief bias in more detail using both the ROC and the forced choice methods.

Chapter 4: Individual Differences

4.1 Introduction

In the previous chapters mixed evidence for response bias and motivated reasoning was found. These results are puzzling when interpreted in light of all accounts of belief bias, including the response bias account. In Experiment 8, cognitive ability was demonstrated to moderate motivated reasoning, showing that only a higher ability subgroup of participants engaged in motivated reasoning. This is an important result because it can explain the discrepancies between Dube, Rotello, and Heit's (2010) findings and ours via a difference in participant sampling. Perhaps Dube et al. sampled a smaller group of above-average cognitive ability participants, masking any potential motivated reasoning effects apparent in the aggregate analysis. The result is also important from a theoretical point of view, because it suggests that any theory of belief bias (and perhaps reasoning in general) explicitly needs to take individual differences into account. Consequently, we proposed an individual differences account of belief bias according to which cognitive ability is a prerequisite to engage in motivated reasoning. Before drawing such far-reaching conclusions, however, a more detailed investigation into the reliability of the effects is required. If these findings turn out to be reliable across methods, a large majority of the current belief bias theories – including the response bias account – will not be able to tell the full story without taking individual differences into account. In the current chapter we present three belief bias experiments using the forced choice and ROC methods investigating the impact of individual differences.

4.1.1 Individual Differences in Reasoning

The impact of individual differences on reasoning has been extensively studied (see Stanovich & West, 2000; 2008). Generally speaking, it is found that the degree of bias in a variety of reasoning tasks is reliably correlated with individual differences in various traits, the main one being cognitive ability – possibly due to its link with working memory capacity (Conway, Kane, & Engle, 2003). These findings are often explained in terms of so-called dual-process theories (DPT) of reasoning. According to DPT, the three main individual differences variables most relevant to explain differential levels of biased responding between subjects are mindware, cognitive ability, and cognitive style (Evans, 2007; 2011; Stanovich & West, 2008).

Mindware refers to the availability of the relevant analytic knowledge required to solve a task, such as the rules of logic or the theorems of probability. Absence of the relevant mindware (often referred to as a mindware gap) makes it impossible for participants to correctly perform the task, making them default to a simpler strategy. An example of a mindware-gap explanation for belief bias can be found in the theory of misinterpreted necessity: according to this theory participants have a poor understanding of "necessity", causing them to use a simpler belief strategy as an escape hatch when application of the necessity concept is required to reason correctly (Evans et al., 1983).

Cognitive ability (commonly referred to as general intelligence or g) is highly correlated to working memory capacity (WMC; Conway et al., 2003). Individual differences in cognitive ability (Newstead, Handley, Harley, Wright, & Farrelly, 2004; Shikishima, Yamagata, Hiraishi, Sugimoto, Murayama, & Ando, 2011) and working memory capacity (Capon, Handley, & Dennis, 2003; Handley, Capon, Beveridge, Dennis, & Evans, 2004) are linked to reasoning performance, with higher ability people generally reasoning better. Consequently, DPTs predict that lower cognitive ability people are generally more biased, mainly due to working memory constraints (e.g., Sá, West, & Stanovich, 1999). An explanation of belief bias in terms of cognitive ability is provided by the metacognitive uncertainty account (Quayle & Ball, 2000). According to this account, certain types of invalid arguments are too difficult because multiple models of the premises need to be simultaneously held in working memory. Only high ability participants have the necessary

WMC available to do this, with others failing and defaulting to beliefs in a similar manner as predicted by the misinterpreted necessity account.

Analytic cognitive style or cognitive reflection (hereafter, cognitive style) refers to the dispositional (Frederick, 2005) *willingness* to engage in analytic processing (e.g., Pennycook, Cheyne, Seli, Koehler, & Fugelsang, 2012). According to DPT, participants with a lower cognitive style disposition have a lower willingness to engage analytic processing. As such, they may rely on simpler strategies leading to more biased responding even if the relevant mindware and cognitive ability levels are available. Cognitive style can be measured using self-report questionnaires such as the rational-experiental inventory (Pacini & Epstein, 1999), the actively open-minded thinking scale (Stanovich & West, 1997), of the need for cognition scale (Cacioppo & Petty, 1982). Alternatively, cognitive style can be measured using a performance measure known as the cognitive reflection test (CRT: Kahneman & Frederick, 2002; Frederick, 2005). This test consists of three questions in which an intuitively compelling initial response should be inhibited in favour of a (relatively simple) correct response. None of the traditional theories of belief bias make explicit predictions about the role of cognitive style, although certain predictions can readily be drawn from a general DPT perspective. For instance, one might predict that participants lower in cognitive style are less likely to engage analytic processing and more likely to rely on a less effortful and more appealing belief-heuristic, leading to increased response bias.

Sá et al. (1999) investigated the impact of individual differences on belief bias and discovered that the response bias component negatively correlated with cognitive ability and self-reported cognitive style, in line with DPT predictions. The researchers did not focus on motivated reasoning – they exclusively compared the conflict (valid-unbelievable and invalid-believable) with the non-conflict (valid-believable and invalid-unbelievable) syllogisms. Furthermore, even if they had analysed the data in the traditional manner, the findings would not have been very indicative of the role of motivated reasoning for two reasons. First, the presented syllogisms were relatively simple (e.g., "all plants need water", "all roses need water", does it follow that all roses are plants?). As we demonstrated in Experiment 5, simple syllogisms do not appear to give rise to motivated reasoning, possibly because of the limited amount of mental models available. Second, premise believability was not controlled for, even though this has been shown to influence belief bias on top of conclusion believability (Thompson, 1996). Finally, the researchers did not use a valid measure of reasoning accuracy (SDT), suggesting that the absence or presence of the logic x belief interaction in a more traditional analysis would not have provided much information about motivated reasoning. Other studies which did investigate the link between motivated reasoning and cognitive ability produced inconsistent results, with some researchers finding that the logic x belief interaction was reduced in a high ability sample (Quayle & Ball, 2000), with others finding that the interaction was reduced in a low ability subgroup (Newstead, Handley, Harley, Wright, & Farrelly, 2004). One possible explanation for these inconsistencies is that none of these experiments adopted the appropriate measurement method.

In the current chapter three studies taking an individual differences approach are presented. The aim of these experiments was to further the motivated reasoning debate, to address the inconsistencies in the literature, and to assess the viability of our newly proposed individual differences account of belief bias. The aim of Experiment 9 was to extend the main finding of Experiment 8, i.e., that motivated reasoning exclusively occurred in a higher cognitive ability subgroup, from the ROC to the forced choice method. In Experiment 10, we investigated the effect of analytic cognitive style as measured with the CRT to see whether the findings generalised to an important alternative individual differences measure. Both of these studies employed the non-simultaneous forced choice procedure. Experiment 11, finally, was specifically designed to compare the relative contributions of cognitive ability and cognitive style to both belief bias components using a large participant sample. This experiment used the ROC method.

4.2 Experiment 9

The aim of this experiment was to find converging evidence for the hypothesis that motivated reasoning is engaged in mainly by a higher cognitive ability subgroup. We employed the non-simultaneous forced choice method, presenting participants with the motivated reasoning (believable v. unbelievable) and response bias (non-conflict v. conflict) problem types. According to the individual differences account of belief bias, cognitive ability is a prerequisite for engaging in motivated reasoning. Thus, the higher ability group should show higher reasoning accuracy for unbelievable compared to believable problems on average, whereas no such difference should occur for the lower ability group. According to the response bias account, motivated reasoning should not occur, regardless of cognitive ability. Instead, all participants should show response bias, resulting in a higher proportion correct for the non-conflict compared to the conflict problem type.

4.2.1 Method

Participants

A total of 108 (17 male) Plymouth University undergraduate psychology students participated in exchange for partial course credit (age range: 18 – 40, *M* = 20, *SD* = 4).

Materials and Measures

A unique list with 64 forced choice reasoning problems was created for each participant. As in the previous forced choice studies, half the problems were designed to measure motivated reasoning (16 believable and 16 unbelievable), the remaining problems were designed to measure response bias (16 non-conflict and 16 conflict).

Cognitive ability was measured using the AH4 test of general intelligence (Heim, 1970). Accuracy, confidence, and response latency on the reasoning task were also measured.

Procedure and Design

The participants were tested in small groups. The cognitive ability test was administered prior to the reasoning task. The non-simultaneous presentation method was employed in order to induce conclusion-to-premise reasoning (cf. Experiment 4).

The design was a 4 (problem type: believable v. unbelievable v. non-conflict v. conflict) x 2 (cognitive ability: higher v. lower) mixed design with the first factor manipulated within subjects and the second manipulated between subjects. Participants who scored at or above the median AH4 score of 90 were assigned to the high cognitive ability group (n = 58), the remaining participants were assigned to the low ability group (n = 50).

4.2.2 Results

Accuracy

Participants were not biased towards the left or the right argument, $t(107) = 1.55$, $p = .12$. We analysed the accuracy data for the motivated reasoning and the response bias items separately. Twelve participants (< 12%) performed below chance for the believable and unbelievable problems and were removed prior to the analyses. As in all the previous experiments, proportion correct was arcsine transformed to conform to the assumptions of ANOVA.

Motivated reasoning

In order to test for motivated reasoning proportion correct was submitted to a 2 (problem type: believable v. unbelievable) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first factor. The higher ability group (M = .78) reasoned better than the lower ability group (M = .68), $F(1, 94) = 15.0$, $p < .001$, $\eta_p^2 = .14$. Cognitive ability and problem type interacted, $F(1, 94) = 9.58$, $p = .003$, $\eta_p^2 = .092$. Follow-up tests revealed that the higher ability group reasoned significantly better for the unbelievable than for the believable problems, $t(50) = 3.02$, $p = .004$. For the lower ability group there was no such difference, $t(44) = 1.41$, $p = .17$. No other effects were significant, all *p*s > .30. Means and standard errors can be found in Table 4.1.

Response bias

To test for response bias, proportion correct was analysed using a 2 (problem type: non-conflict v. conflict) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first factor. The analysis revealed a main effect of problem type showing that accuracy was higher for the non-conflict ($M = .74$) than for the conflict ($M = .67$) problem types, $F(1, 94) = 7.7$, $p = .007$, $\eta_p^2 = .075$. Accuracy was also higher for the higher ($M = .76$) than for the lower ($M = .65$) cognitive ability group, $F(1, 94) = 17.2$, $p < .001$, $\eta_p^2 = .16$. Problem type and cognitive ability marginally interacted, $F(1, 94) = 3.7$, $p = .059$, $\eta_p^2 = .037$. Follow-up tests revealed that for the higher ability group there was no difference in accuracy for the non-conflict and the conflict problems, $t(50) = 0.63$, $p = .53$. For the lower ability group, proportion correct was significantly higher for the non-conflict than for the conflict problems, $t(44) = 3.18$, $p = .003$. Means and standard errors can be found in Table 4.1.

Table 4.1

Experiment 9: Means (Standard Errors) for the Accuracy, Confidence Rating and Response Time Analysis

Variable	Group	Believable	Unbelievable	Non-conflict	Conflict
Accuracy	**Higher ability**	.76 (.02)	.81 (.02)	.77 (.02)	.75 (.02)
	Lower ability	.70 (.02)	.66 (.02)	.71 (.02)	.59 (.03)
Confidence	**Higher ability**	2.13 (0.07)	2.19 (0.07)	2.19 (0.07)	2.24 (0.07)
	Lower ability	1.99 (0.07)	1.97 (0.07)	2.02 (0.07)	2.00 (0.07)
Latency	**Higher ability**	20,416 (831)	19,279 (816)	19,828 (904)	18,941 (807)
	Lower ability	20,397 (831)	19,698 (816)	18,225 (904)	18,587 (872)

Note. Accuracy is proportion correct responses, confidence is on a scale from 1 – 3, response time is measured in ms.

Confidence

Motivated reasoning

Confidence ratings were submitted to a 2 (problem type: believable v. unbelievable) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first factor. Higher ability participants were marginally more confident (M = 2.16) than the lower ability participants (M = 1.98), $F(1, 94) = 3.41$, $p = .068$, $\eta_p^2 = .035$. Problem type and cognitive ability marginally interacted, suggesting that the higher ability participants tended to be more confident for the unbelievable than for the believable problems, with the lower ability participants showing the opposite pattern, $F(1, 94) = 2.86$, $p = .094$, $\eta_p^2 = .030$. The remaining effect did not approach significance, $p > .39$. Means and standard errors can be found in Table 4.1.

Response bias

Confidence ratings were submitted to a 2 (problem type: non-conflict v. conflict) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first factor. The higher ability participants were more confident (M = 2.21) than the lower ability participants (M = 2.01), $F(1, 94) = 4.77$, $p = .031$, $\eta_p^2 = .048$. No other effects approached significance, all *p*s > .18. Means and standard errors can be found in Table 4.1.

Latency

Motivated reasoning

Log-transformed response times were submitted to a 2 (problem type: believable v. unbelievable) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first factor. Participants responded more quickly to unbelievable than to believable problems, $F(1, 94) = 4.82$, $p = .031$, $\eta_p^2 = .049$. No other effects were significant, all *p*s > .81. Means and standard errors can be found in Table 4.1.

Response bias

Log-transformed response times were submitted to a 2 (problem type: non-conflict v. conflict) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first factor. No effects approached significance, all *p*s > .18. Means and standard errors can be found in Table 4.1.

4.2.3 Discussion

The aim of the current experiment was to extend the finding that motivated reasoning is linked to cognitive ability to the forced choice paradigm. We found that higher cognitive ability participants engaged in motivated reasoning, whereas no such effect occurred in the lower ability group. This finding echoes the findings in Experiment 8 and suggests that cognitive ability is linked to motivated reasoning. This is more compatible with data presented by Newstead et al. (2004), who exclusively found the logic x belief interaction for their higher ability participants. The findings are incompatible with Quayle and Ball's (2000), who found the opposite – i.e., a larger logic x belief interaction in their lower ability participants. We also investigated response bias. Response bias was exhibited by all the participants, although there was a tendency for the effect to be driven mainly by the low ability subgroup. This finding is compatible Sá et al.'s (1999) findings.

These findings are inconsistent with the response bias account of belief bias, which suggests that motivated reasoning is not a component of belief bias. The results also pose problems for many of the traditional belief bias theories, if only for the fact that they do not explicitly consider the impact of individual differences. Two notable exceptions to this are the modified selective processing theory (Stupple et al., 2011) and MMT (Oakhill et al., 1989). According to the former theory, different subgroups of reasoners adopt qualitatively distinct reasoning strategies, as evidenced by different response time patterns. This theory lacks a formal specification of which individual differences variables are linked to subgroup membership, however. Instead, it uses overall reasoning aptitude as the main variable to determine group membership, which is a circular definition. MMT predicts that working memory capacity mediates the search for counterexamples, with those higher in WMC being more likely to be successful at the model search. The absence of motivated reasoning for the lower ability group demonstrated here is compatible with this prediction. The findings are also compatible with the individual differences account of belief bias: the current findings suggest that cognitive ability predicts group membership.

One notable caveat of our findings so far is that the effects are small, suggesting that other factors may also play a role in explaining the degree to which people engage in motivated reasoning. DPT suggests that cognitive style might also play an important role in understanding the link between individual differences and both belief bias components. The role of cognitive style was investigated in Experiment 10.

4.3 Experiment 10

While cognitive ability is a measure of the capability of participants to use analytic processing, cognitive style is a measure of the willingness to do so (e.g., Pennycook et al., 2012). We have demonstrated using both the ROC (Experiment 8) and the forced choice (Experiment 9) methods that only a higher ability subgroup of participants engaged in motivated reasoning. One straightforward explanation for this finding is that lower ability reasoners lack the necessary working memory capacity to engage in the effortful (Deutsch, Kordts-Freudinger, Gawronski, & Strack, 2009) negation process that belief bias accounts such as selective processing (Evans, Handley, & Harper, 2001; Klauer, Musch, & Naumer, 2000) and mental model theory (Oakhill et al., 1989) argue is provoked by unbelievable conclusions. Instead, these participants possibly adopt a confirmatory testing strategy regardless of believability. For these participants, beliefs only affect the decision process in terms of a heuristic response bias. It might appear counterintuitive that higher capacity is linked to increased rather than decreased levels of bias (i.e., motivated reasoning). However, the high ability group compensates by better overall reasoning performance and reduced levels of response bias. Nevertheless, the link between cognitive ability and motivated reasoning does raise the question whether perhaps analytic cognitive style also plays a major role in explaining response bias and motivated reasoning, given its large correlation with cognitive ability.

Research has shown that cognitive ability and cognitive style are moderately to highly correlated, depending on whether a self-report [r(AOT, CA) ≈ .20 – Stanovich & West, 2000] or a performance measure is used [r(CRT,CA) ≈ .50 – Frederick, 2005]. An alternative interpretation for the current findings is that perhaps cognitive ability is linked to motivated reasoning due to its correlation with cognitive style. A first step in investigating this possibility is to figure out whether cognitive style has any bearing on belief bias whatsoever.

In the current experiment we used the non-simultaneous forced choice reasoning method to investigate whether cognitive style predicts motivated reasoning. A secondary aim of the current experiment was to see if the previous findings of motivated reasoning and response bias could be replicated using a between subjects design. In all non-simultaneous forced choice experiments so far, problem type was manipulated within subjects, meaning that all participants solved both the motivated reasoning and response bias problems types. Stanovich and West (2008) argued that a true cognitive bias should occur both in a within- as well as a between-subjects design, because the prior method could lead to unwanted effects (e.g., experimenter effects or an increased impact due to the obvious contrast between the item types – see also the belief suppression hypothesis discussed in Chapter 2). In the present study we used a blocked design to control for this potential problem.

According to the response bias account, participants should not engage in motivated reasoning. On the basis of our previous findings we might predict that motivated reasoning should be engaged in exclusively by a high cognitive style group, although no belief bias accounts explicitly predict this. An absence of motivated reasoning suggests that cognitive style is not a relevant predictor of motivated reasoning and the cognitive ability is the more important factor, which would be compatible with MMT's predictions that WMC is linked to the motivated search for counterexamples.

4.3.1 Method

Participants

A total of 71 undergraduate psychology students (8 male) from Plymouth University volunteered to take part in the experiment in exchange for course credit (age: range = 18 – 54, *M* = 23, *SD* = 7).

Materials

Materials were constructed in the same way as in Experiment 9. All participants were presented with motivated reasoning problems (i.e., believable and unbelievable) and response bias problems (i.e., non-conflict and conflict) for a total of 64 trials.

Cognitive style was measured using the CRT (Kahneman & Frederick, 2002), a behavioural measure of analytic cognitive style (Frederick, 2005). The CRT is a three-question test which is a short and efficient performance measure of analytic cognitive style (e.g., Pennycook et al., 2012; Shenhav, Rand, & Greene, 2011). The participant has to solve three questions which cue an intuitively compelling (but incorrect) response. The three questions are:

1. A bat and a ball cost $1.10. The bat costs $1 more than the ball. How much does the ball cost? ____ cents.

2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? ____ minutes.

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

For the bat-ball problem, the correct response is 5 cents, because the bat costs $1 more than the ball (i.e., the bat costs $1.05 and the ball costs $0.05). Due to attribute substitution, however, participants who lack the motivation or willingness to engage in the fairly low degree of effortful reasoning required to give the correct response will incorrectly respond 10 cents. For the widget problem, the correct response equals five minutes because it takes one machine five minutes to make a widget. The intuitive response is 100 minutes due to matching bias. Finally, for the lily pad problem, the correct response is 47, because the relation between time and lake coverage is exponential. Incorrectly assuming the relation is linear will lead to the incorrect response of 24 days.

Procedure, Measures, and Design

The procedure was very similar to that of the previous experiment, with the exception that upon completing the experiment, the participants were asked to complete the CRT. We measured proportion correct responses, confidence ratings, response latency, and CRT score (0 – 3).

The design differed somewhat from the previous experiments. Participants were presented with the motivated reasoning and response bias problem types in a blocked manner. Half the participants were randomly assigned to a motivated reasoning first condition, in which the first 32 problems were of the believable and unbelievable types and the final 32 problems of the non-conflict and conflict types. The remaining participants were presented with the response bias problems first, followed by the motivated reasoning problems. Participants were not explicitly made aware of the blocked nature of this design: there was no pause in between the blocks, nor were there any explicit instructions hinting at the different blocks. The experiment employed a mixed 4 (problem type: believable v. unbelievable v. non-conflict v. conflict) x 2 (analytic cognitive style: lower v. higher) x 2 (block order: motivated reasoning first v. response bias first) design. Participants were assigned to the lower cognitive style group if they did not correctly solve any of the CRT questions (n = 58). The remaining participants were assigned to a higher cognitive style condition (n = 29).

4.3.2 Results

Accuracy

Analysis comparing only the first block demonstrated that the results of the between-subjects and the full analysis did not differ, thus only the latter is reported here. The participants were not spatially biased, $t(70) = 1.37$, $p = .17$. We analysed the motivated reasoning and response bias conditions separately. Nine participants (< 13%) responded below chance for the believable and unbelievable problems and were removed prior to the analysis. All proportions were arcsine transformed prior to the analyses.

Motivated reasoning

In order to test for motivated reasoning, proportions correct were submitted to a 2 (problem type: believable v. unbelievable) x 2 (cognitive style: higher v. lower) x 2 (block order: motivated reasoning problems first v. response bias problems first) mixed ANOVA with repeated measures on the first factor. Problem type and cognitive style significantly interacted, $F(1, 58) = 12.48$, $p = .001$, $\eta_p^2 = .18$. Follow-up tests revealed that the higher cognitive style group reasoned better for the unbelievable ($M = .80$) than for the believable ($M = .72$) problems, $t(26) = 3.03$, $p = .005$. For the lower style group, this difference was not significant, $t(34) = 1.34$, $p = .19$. No other effects reached significance, all *p*s > .10. Means and standard errors can be found in Table 4.2.

Response bias

In order to test for response bias, proportions correct were submitted to a 2 (problem type: non-conflict v. conflict) x 2 (cognitive style: higher v. lower) x 2 (block order: motivated reasoning problems first v. response bias problems first) mixed ANOVA with repeated measures on the first factor. Accuracy was higher for the non-conflict ($M = .77$) than for the conflict ($M = .64$) problems, $F(1, 58) = 11.17$, $p = .001$, $\eta_p^2 = .16$. The higher style subgroup had higher accuracy overall ($M = .75$) than the lower style group ($M = .65$), $F(1, 58) = 6.80$, $p = .012$, $\eta_p^2 = .11$. Participants given the response bias block first had higher accuracy ($M = .73$) than those who solved the motivated reasoning block first ($M = .68$), $F(1, 58) = 6.80$, $p = .012$, $\eta_p^2 = .11$. Problem type and cognitive style interacted, $F(1, 58) = 4.73$, $p = .034$, $\eta_p^2 = .075$. Follow-up tests revealed that for the higher cognitive style subgroup there was no accuracy difference between the non-conflict and conflict problems, $t(26) < 1$, $p = .33$. The lower style subgroup performed better for the non-conflict ($M = .76$) than for the conflict ($M = .54$) problems, $t(34) = 3.36$, $p = .002$. Problem type also interacted with block order, $F(1, 58) = 4.47$, $p = .039$, $\eta_p^2 = .072$. Follow-up tests revealed that in the response bias first block accuracy was significantly higher for the non-conflict ($M = .83$) than for the conflict ($M = .63$) problems, $t(30) = 3.14$, $p = .004$. In the motivated reasoning first block this difference was not significant, $t(30) = 1.29$, $p = .21$. No other effects were significant, all $ps > .17$. Means and standard errors can be found in Table 4.2.

Table 4.2

Experiment 10: Means (Standard Errors) for the Accuracy, Confidence Rating and Response Time Analysis

Var	Block	Style	Believable	Unbelievable	Non-conflict	Conflict
Acc	**MR first**	**High**	.69 (.05)	.77 (.06)	.73 (.05)	.67 (.07)
		Low	.67 (.03)	.64 (.04)	.65 (.03)	.54 (.05)
	RB first	**High**	.71 (.04)	.77 (.05)	.77 (.04)	.74 (.06)
		Low	.69 (.04)	.59 (.05)	.84 (.04)	.46 (.06)
Conf	**MR first**	**High**	2.19 (0.14)	2.18 (0.14)	2.10 (0.14)	2.08 (0.14)
		Low	1.95 (0.09)	1.92 (0.09)	1.89 (0.09)	1.81 (0.09)
	RB first	**High**	1.94 (0.11)	2.01 (0.11)	2.16 (0.11)	2.27 (0.11)
		Low	1.84 (0.11)	1.64 (0.12)	2.22 (0.12)	1.95 (0.11)
RT	**MR first**	**High**	23,244 (2,098)	21,576 (1,944)	16,282 (1,544)	17,555 (1,830)
		Low	20,851 (1,365)	21,220 (1,265)	13,041 (1,004)	13,581 (1,190)
	RB first	**High**	15,333 (1,640)	14,812 (1,520)	20,511 (1,208)	20,311 (1,431)
		Low	15,935 (1,740)	16,016 (1,612)	18,428 (1,280)	19,289 (1,518)

Note. Accuracy is proportion correct responses, confidence is on a scale from 1 – 3, and response time is measured in ms.

Confidence

Motivated reasoning

Confidence ratings were analysed using a 2 (problem type: believable v. unbelievable) x 2 (cognitive style: higher v. lower) x 2 (block order: motivated reasoning block first v. response bias block first) mixed ANOVA with repeated measures on the first factor. A significant main effect of block order was found, indicating that confidence was higher for those who were presented with the motivated reasoning problems first (M = 2.11) compared to those presented with the response bias problems first (M = 1.86), $F(1, 58)$ = 4.97, p = .030, η_p^2 = .079. Confidence was also marginally larger for the higher cognitive style (M = 2.09) compared to the lower cognitive style (M = 1.88) group, $F(1, 58)$ = 3.44, p = .069, η_p^2 = .056. Cognitive style marginally interacted with problem type, suggesting that the higher style subgroup was more confident for the unbelievable than for the believable problems, with the opposite pattern emerging for the lower style participants, $F(1, 58)$ = 3.31, p = .074, η_p^2 = .054. No other effects approached significance, all ps > .13. Means and standard errors can be found in Table 4.2.

Response bias

Confidence ratings were analysed using a 2 (problem type: non-conflict v. conflict) x 2 (cognitive style: higher v. lower) x 2 (block order: motivated reasoning block first v. response bias block first) mixed ANOVA with repeated measures on the first factor. Confidence was higher for the non-conflict (M = 2.09) than for the conflict (M = 2.03) problems, $F(1, 58)$ = 4.25, p = .044, η_p^2 = .068. Problem type and cognitive style interacted, $F(1, 58)$ = 8.91, p = .004, η_p^2 = 13. Follow-up tests revealed that those lower in cognitive style were more confident for the non-conflict (M = 2.07) than for the conflict (M = 1.91) problems, $t(34)$ = 3.52, p = .001. There was no such difference for those higher in cognitive style, $t(26)$ = 1.11, p = .28. Problem type, cognitive style, and block order also marginally interacted, $F(1, 58)$ = 3.12, p = .083, η_p^2 = .051. This effect suggested that the lower cognitive style participants who were presented with the response bias block first had a large drop in confidence: they were highly confident for the non-conflict problems (*M* = 2.22), but much less confident for the unbelievable problems (*M* = 1.64). No other effects approached significance, all *p*s > .21. Means and standard errors can be found in Table 4.2.

Latency

Motivated reasoning

Log-transformed response latency was analysed using a 2 (problem type: believable v. unbelievable) x 2 (cognitive style: higher v. lower) x 2 (block order: motivated reasoning block first v. response bias block first) mixed ANOVA with repeated measures on the first factor. The analysis only resulted in a main effect of block order, indicating that those given the motivated reasoning problems first took much longer (*M* = 22,869 ms) to respond than those given the response bias problems first (*M* = 15,669 ms), $F(1, 58) = 14.46$, $p < .001$, $\eta_p^2 = .20$. No other effects approached significance, all *p*s > .19. Means and standard errors can be found in Table 4.2.

Response bias

Log-transformed response latency was analysed using a 2 (problem type: non-conflict v. conflict) x 2 (cognitive style: higher v. lower) x 2 (block order: motivated reasoning block first v. response bias block first) mixed ANOVA with repeated measures on the first factor. Participants given the response bias problems first took longer to respond (*M* = 19,883 ms) than those given the motivated reasoning problems first (*M* = 15,560 ms), $F(1, 58) = 8.28$, $p = .006$, $\eta_p^2 = .13$. There was also a marginal main effect of cognitive style, suggesting that those higher in cognitive style took longer to respond (*M* = 18,837 ms) than those of lower cognitive style (*M* = 16,605 ms), $F(1, 58) = 3.42$, $p = .07$, $\eta_p^2 = .056$. No other effects approached significance, all *p*s > .16. Means and standard errors can be found in Table 4.2.

4.3.3 Discussion

The aim of Experiment 10 was to investigate whether analytic cognitive style was linked to motivated reasoning. The results indicated that motivated reasoning occurred mainly in a higher cognitive style subgroup. Response bias was found for all participants regardless of cognitive ability, although the magnitude of the effect was larger for those of lower cognitive style, and for those presented with the response bias problems first. These findings also demonstrate that the forced choice paradigm can be used to measure both components of belief bias using a between or a within subjects manipulation, as long as individual differences are controlled for.

A clear pattern on the relation between motivated reasoning and individual differences has emerged: using the ROC and the forced choice procedure we demonstrated that those of higher cognitive ability and/or style were more likely to engage in motivated reasoning than those of relatively lower ability/style. This naturally raises the question which individual difference is the most potent predictor: is adequate cognitive ability required in order to be able to engage in motivated reasoning, or is cognitive style the necessary prerequisite for motivated reasoning? We addressed this issue in our final experiment.

4.4 Experiment 11

The aim of this final experiment was to compare the relative importance of cognitive ability and cognitive style as determinants of both belief bias components. Having established that cognitive style and cognitive ability both predict motivated reasoning, we investigated whether ability or style was the crucial factor. Our use of the ROC method also allowed us to investigate whether the cognitive style effect could be replicated using a different method.

We tested a large participant sample using the belief bias confidence rating task as used in Experiments 5 – 8. Our findings so far suggest that neither the response bias account nor the traditional accounts appear to provide a viable interpretation of the presented belief bias patterns, although, as predicted by all of these theories, response bias still remains a major component. Based on our previous findings, we predict that both higher cognitive ability and higher cognitive style are linked to motivated reasoning, with response bias occurring more generally for all participants, albeit in a greater degree for the lower ability and/or style participants. The main analysis (excluding the endorsement and latency analyses) was a three-step process. First, accuracy and bias SDT parameters were compared as a function of cognitive ability and cognitive style in separate analyses. Next, mediational analysis was used to investigate whether style or ability respectively mediated the link between ability or style and motivated reasoning. The same mediational analyses were also performed for response bias. Finally, a path model was created to account of the effect of individual differences in cognitive ability and cognitive style on motivated reasoning, response bias, response time taken, and reasoning performance.

4.4.1 Method

Participants

A total of 191 University of Waterloo (Canada) undergraduates volunteered to take part in the study (62 male, 129 female, age: range = 17 – 50, M = 20, SD = 3).

Materials and Measures

The materials consisted of 64 complex (multiple model) syllogisms, half of which were valid and half of which were invalid. Item contents were randomly assigned to the syllogistic problem frames for each participant anew as in Experiments 5 – 8.

The cognitive ability and cognitive style measures differed somewhat from the previous experiments. Participants were given six different measures that have been used in past research to differentially measure cognitive ability or cognitive style (e.g., Pennycook et al., 2012). The key factor that distinguishes these measures is the presence of a misleading intuitive response cue. The cognitive ability and cognitive style measures were roughly equivalent in terms of difficulty, but the former measures did not cue misleading intuitive responses. Consequently, in order to get a higher cognitive ability score the participant needed to engage analytic processing. In contrast, for participants to get a higher analytic cognitive style score, a compelling intuitive response needed to be resisted in favour of some relatively light analytic processing. For the cognitive style measure participants completed 3 CRT problems (as in Experiment 10), 6 incongruent base-rate problems (De Neys & Glumicic, 2008), and 18 ratio bias problems (Bonner & Newell, 2010). For the cognitive ability measure participants completed 3 numeracy problems (Schwartz, Woloshin, Black, & Welch, 1997), 6 neutral base-rate problems (De Neys & Glumicic, 2008), and the 10 item WordSum verbal intelligence test (Huang & Hauser, 1998). These problems can be found in Appendix A. Cognitive ability and cognitive style scores were computed by averaging across the mean accuracy of the measures outlined above.

Procedure and Design

The procedure was identical to that of Experiments 5 – 8 (Chapter 3), with the exception that cognitive ability and cognitive style were measured prior to the syllogistic reasoning task. The individual differences tasks were administered in the following order: base-rate (neutral and conflict), ratio-bias, CRT, numeracy, WordSum.

The experiment used a 2 (logical validity: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (analytic ability/style: higher v. lower) mixed design, with the first two factors manipulated within subjects and the ability and style variables manipulated between subjects. Participants were assigned to a higher or lower ability and style group based on whether they scored at or above v. below the median (analytic style: *Mdn* = .625; cognitive ability: *Mdn* = .75). Note that ability and style were not crossed in the first step of this analysis. One participant was removed from all further analyses because we failed to collect individual differences data.

4.4.2 Results

Endorsement

Endorsement rates were analysed to allow for a comparison with traditional belief bias results reported in the literature. We analysed the results for ability and style separately.

Cognitive ability

Endorsement rates were submitted to a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first two factors. Valid arguments were endorsed more than invalid arguments, $F(1, 188) = 201.08$, $p < .001$, $\eta_p^2 = .52$. Believable arguments were endorsed more than unbelievable arguments, $F(1, 188) = 106.14$, $p < .001$, $\eta_p^2 = .36$. The two factors also significantly interacted indicating that the validity effect was larger for unbelievable than for believable problems, $F(1, 188) = 5.54$, $p = .020$, $\eta_p^2 = .029$. Cognitive ability interacted with validity showing that the validity effect was larger for the higher ability subgroup, $F(1, 188) = 17.35$, $p < .001$, $\eta_p^2 = .084$. Ability also interacted with believability suggesting that the belief effect was larger for the lower ability subgroup, $F(1, 188) = 6.79$, $p = .010$, $\eta_p^2 = .035$. Finally, the three factors also significantly interacted, $F(1, 188) = 9.37$, $p = .003$, $\eta_p^2 = .047$. Follow-up tests for the higher and lower groups separately indicated that for the high ability subgroup a significant logic x belief interaction occurred, $F(1, 94) = 22.93$, $p < .001$, $\eta_p^2 = .20$. The logic x belief interaction was not significant for the low ability group, $F(1, 94) < 1$, $p = .67$. The ability main effect did not approach significance, $p = .62$. Means and standard errors can be found in Table 4.3.

Cognitive style

Endorsement rates were submitted to a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (cognitive style: higher v. lower) mixed ANOVA with repeated measures on the first two factors. Valid arguments were endorsed more than invalid ones, $F(1, 188) = 196.78$, $p < .001$, $\eta_p^2 = .51$. Believable arguments were endorsed more than unbelievable ones, $F(1, 188) = 121.39$, $p < .001$, $\eta_p{}^2 = .39$. Logic and belief interacted in such a way that the validity effect was larger in the unbelievable condition, $F(1, 188) = 4.75$, $p = .031$, $\eta_p{}^2 = .025$. Cognitive style interacted with validity showing that the higher style subgroup was better at discriminating between valid and invalid arguments compared to the lower style group, $F(1, 188) = 20.40$, $p < .001$, $\eta_p{}^2 = .098$. Style interacted with believability in such a way that the lower style subgroup was more influenced by their prior beliefs than the high style group, $F(1, 188) = 24.86$, $p < .001$, $\eta_p{}^2 = .12$. Finally, the three-way interaction between logic, belief and style was also significant, $F(1, 188) = 15.78$, $p < .001$, $\eta_p{}^2 = .077$. Follow up analyses revealed that the logic x belief interaction was significant in the higher style group, $F(1, 99) = 22.93$, $p < .001$, $\eta_p{}^2 = .18$, but not in the lower style group, $F(1, 89) = 1.39$, $p = .24$, $\eta_p{}^2 = .015$. The main effect of style did not approach significance, $p = .78$. Means and standard errors can be found in Table 4.3.

Table 4.3

Experiment 11: Means (Standard Errors) for the Endorsement Rate Analysis for Cognitive Ability and Cognitive Style

	Valid		**Invalid**	
	Believable	**Unbelievable**	**Believable**	**Unbelievable**
Higher Ability	.84 (.02)	.71 (.03)	.57 (.03)	.30 (.03)
Lower Ability	.85 (.02)	.53 (.03)	.68 (.03)	.35 (.03)
Higher Style	.82 (.02)	.74 (.03)	.56 (.03)	.32 (.02)
Lower Style	.87 (.02)	.49 (.03)	.70 (.03)	.34 (.03)

Note. Ability = cognitive ability, style = analytic cognitive style.

SDT

As in Experiments 5 – 8 we plotted ROCs (Figure 4.1) to which we fit SDT models for each participant to estimate accuracy (A_z) and bias (c_a). The parameters were analysed as a function of cognitive ability and cognitive style separately.

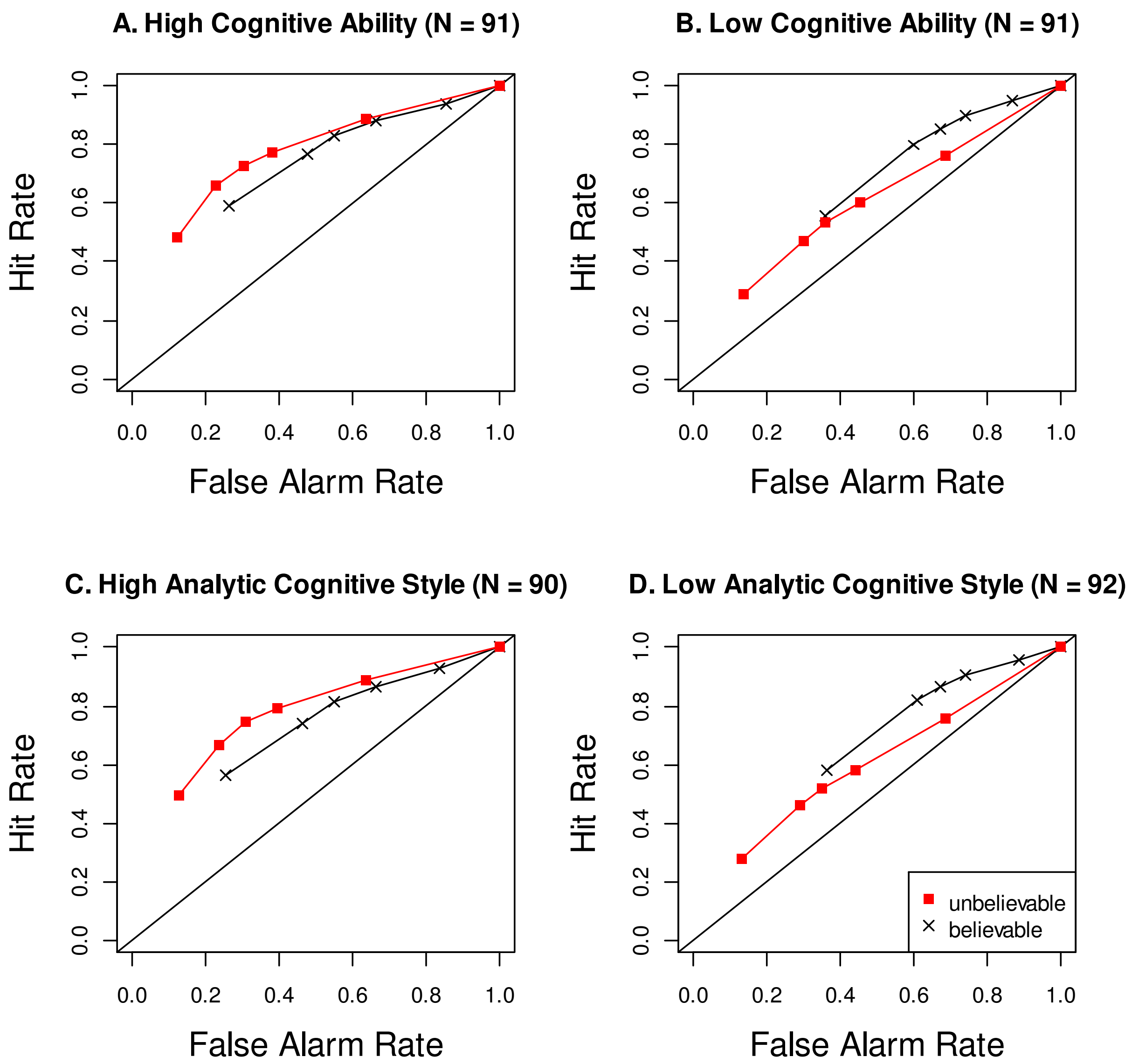


Figure 4.1. Top row: Aggregate ROCs for the higher (A) and lower (B) cognitive ability groups. Bottom row: Aggregate ROCs for the higher (C) and lower (D) cognitive ability groups.

Cognitive ability

Accuracy was analysed using a 2 (conclusion believability: believable v. unbelievable) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first factor. Participants reasoned better for unbelievable (A_z = .68) than for believable (A_z = .66) syllogisms, $F(1, 188) = 4.49$, $p = .035$, $\eta_p^2 = .023$. Higher ability participants (A_z = .72) also reasoned better than lower ability ones (A_z = .61), $F(1, 188) = 26.48$, $p < .001$, $\eta_p^2 = .12$. Belief and ability interacted, $F(1, 188) = 8.91$, $p = .003$, $\eta_p^2 = .045$. Follow up tests revealed that the higher ability group reasoned better for unbelievable than for believable problems, $t(94) = 4.38$, $p < .001$. This difference was absent in the lower ability group, $t(94) < 1$, $p = .60$. Means and standard errors can be found in Table 4.4.

Response bias was analysed using a 2 (conclusion believability: believable v. unbelievable) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first factor. Participants adopted a more liberal response criterion for the believable (c_a = -0.68) than for the unbelievable (c_a = 0.05) problems, $F(1, 188) = 108.23$, $p < .001$, $\eta_p^2 = .37$. Ability and believability interacted, suggesting that the belief effect was larger in the lower ability subgroup, $F(1, 188) = 5.25$, $p = .023$, $\eta_p^2 = .027$. Follow-up tests revealed that the belief effect was significant for those of higher, $t(94) = 9.02$, $p < .001$ and lower, $t(94) = 5.71$, $p < .001$ cognitive ability. There was no main effect of cognitive ability, $p = .87$. Means and standard errors can be found in Table 4.4.

Table 4.4

Experiment 11: Means (Standard Errors) for the SDT Analysis for Cognitive Ability and Cognitive Style

	Accuracy (A_z)		**Bias (c_a)**	
	Believable	**Unbelievable**	**Believable**	**Unbelievable**
Higher Ability	.69 (.02)	.75 (.02)	-0.61 (0.06)	-0.04 (0.07)
Lower Ability	.62 (.02)	.60 (.02)	-0.76 (0.06)	0.13 (0.07)
Higher Style	.69 (.02)	.76 (.02)	-0.53 (0.06)	-0.10 (0.07)
Lower Style	.62 (.02)	.59 (.02)	-0.84 (0.06)	0.21 (0.07)

Note. Lower bias values indicate more liberal criterion placement.

Cognitive style

Accuracy was analysed using a 2 (conclusion believability: believable v. unbelievable) x 2 (cognitive style: higher v. lower) mixed ANOVA with repeated measures on the first factor. There was a marginal main effect of belief, suggesting that reasoning accuracy tended to be higher for the unbelievable (A_z = .67) than for the believable (A_z = .65) problems, $F(1, 188) = 3.79$, $p = .053$, $\eta_p^2 = .020$. Higher style participants also reasoned better (A_z = .72) than lower style ones (A_z = .60), $F(1, 188) = 28.57$, $p < .001$, $\eta_p^2 = .13$. Belief and style interacted, $F(1, 188) = 15.02$, $p < .001$, $\eta_p^2 = .074$. Follow up tests revealed that the higher cognitive style group reasoned better for the unbelievable than the believable arguments, $t(99) = 4.78$, $p < .001$. For the low style group, no such difference occurred, $t(89) = 1.19$, $p = .24$. Means and standard errors can be found in Table 4.4.

Bias was analysed using a 2 (conclusion believability: believable v. unbelievable) x 2 (cognitive style: higher v. lower) mixed ANOVA with repeated measures on the first factor. Participants adopted a more liberal response criterion for the believable (c_a = -0.69) compared to the unbelievable (c_a = 0.06) problems, $F(1, 188) = 122.43$, $p < .001$, $\eta_p^2 = .39$. Belief and style interacted suggesting that the belief effect is larger for the lower compared to the higher style group, $F(1, 188) = 21.63$, $p < .001$, $\eta_p^2 = .10$. Follow up tests confirmed that the belief effect was significant for both the higher, $t(99) = 9.44$, $p < .001$ and lower group, $t(89) = 5.51$, $p < .001$. There was no main effect of style, $p = .97$. Means and standard errors can be found in Table 4.4.

Latency

Log transformed response times were analysed separately for cognitive ability and cognitive style.

Cognitive ability

We analysed RT using a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (cognitive ability: higher v. lower) mixed ANOVA with repeated measures on the first two factors. Participants responded more quickly to valid (M = 13,427 ms) than to invalid (M = 16,487 ms) problems, $F(1, 188) = 104.30$, $p < .001$, $\eta_p^2 = .36$. Higher ability participants responded significantly more slowly (M = 16,814 ms) than lower ability (M = 13,099 ms) ones, $F(1, 188) = 8.12$, $p = .005$, $\eta_p^2 = .041$. Logic and ability interacted, suggesting that the validity effect was larger for the higher ability subgroup than for the lower ability group, $F(1, 188) = 6.64$, $p = .011$, $\eta_p^2 = .034$. Follow up tests confirmed that the validity effect was significant for both the higher, $t(94) = 8.25$, $p < .001$, and lower ability group, $t(94) = 5.64$, $p < .001$. No other effects were significant, all *p*s > .11. Means and standard errors can be found in Table 4.5.

Cognitive style

RTs were analysed using a 2 (logical status: valid v. invalid) x 2 (conclusion believability: believable v. unbelievable) x 2 (cognitive style: higher v. lower) mixed ANOVA with repeated measures on the first two factors. Participants responded more quickly to valid (M = 13,325 ms) than to invalid (M = 16,331 ms) problems, $F(1, 188) = 102.37$, $p < .001$, $\eta_p^2 = .35$. Participants higher in cognitive style also responded more slowly (M = 17,273 ms) than those lower in style (M = 12,383 ms), $F(1, 188) = 16.93$, $p < .001$, $\eta_p^2 = .083$. Logic and style interacted, $F(1, 188) = 9.88$, $p = .002$, $\eta_p^2 = .050$. Follow-up tests revealed that the validity effect was significant for both the higher cognitive style group, $t(99) = 9.29$, $p < .001$, and for the lower cognitive style group, $t(89) = 4.70$, $p < .001$. No other effects were significant, all *p*s > .17. Means and standard errors can be found in Table 4.5.

Table 4.5

Experiment 11: Means (Standard Errors) for the Latency Analysis for Cognitive Ability and Cognitive Style

	Valid		Invalid	
	Believable	**Unbelievable**	**Believable**	**Unbelievable**
Higher Ability	15,199 (707)	14,380 (615)	19,337 (933)	18,339 (904)
Lower Ability	12,070 (707)	12,056 (615)	14,627 (933)	14,002 (904)
Higher Style	15,566 (676)	14,946 (583)	19,594 (897)	18,986 (858)
Lower Style	11,490 (713)	11,299 (614)	13,701 (946)	13,042 (905)

Note. Response latency was measured in milliseconds.

Mediation Analysis

We conducted several mediational analyses in order to answer our main question whether cognitive ability or cognitive style is the more potent predictor of motivated reasoning. We conducted the analyses for motivated reasoning and response bias separately. We calculated a motivated reasoning index (MRI) by subtracting A_z–believable from A_z–unbelievable, with higher values indicating more motivated reasoning. We also calculated a response bias index (RBI) by subtracting c_a–believable from c_a–unbelievable, after adding a constant k (k = 10000, arbitrarily chosen to be large enough) to both values. This constant was added to deal with the fact that for some participants the sign for c_a-believable and unbelievable was not identical, leading to skewed bias estimates. Higher values of RBI indicated more response bias. Prior to conducting the following analyses, 10 outliers were removed using boxplots (Figure 4.2).

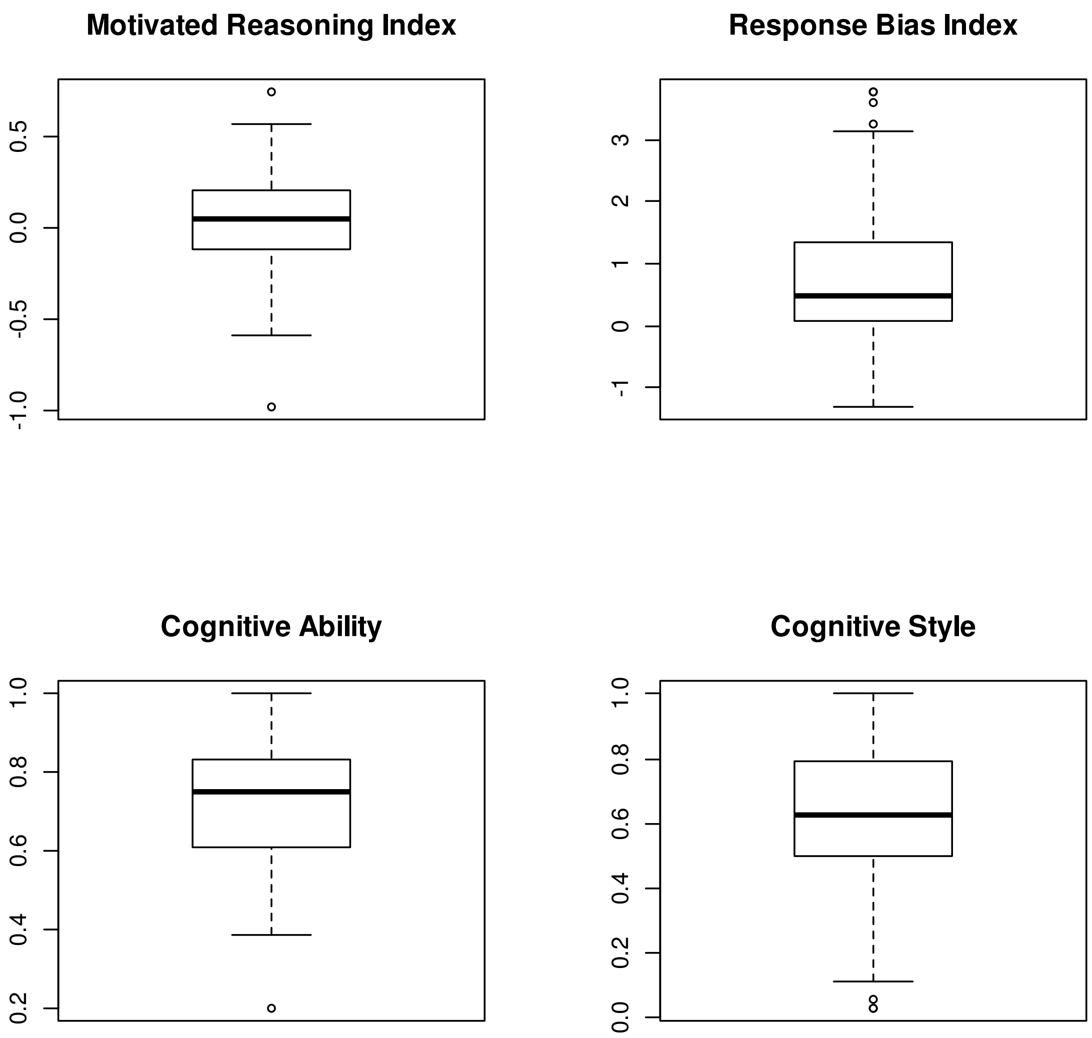


Figure 4.2. Outliers for the variables included in the mediation analyses were determined using boxplots. Including the outliers in the analyses did not change the conclusions. Two outliers overlapped.

Motivated reasoning

We tested whether cognitive ability mediated the link between cognitive style and motivated reasoning. First, we regressed MRI on cognitive style. The analysis revealed that style significantly predicted motivated reasoning, $b = 0.30$, $t(178) = 3.26$, $p = .001$. Second, we added cognitive ability into the regression. The analysis showed that style still significantly predicted motivated reasoning, $b = 0.29$, $t(177) = 2.56$, $p = .011$. In contrast, ability did not significantly predict motivated reasoning, $b = 0.03$, $t(177) < 1$, $p = .84$.

Cognitive ability did not mediate the relationship between cognitive style and cognitive ability, Sobel's $Z = 0.21$, $p = .83$.

We also tested the opposite hypothesis that cognitive style mediated the link between cognitive ability and motivated reasoning. MRI was regressed on cognitive ability, indicating that ability significantly predicted motivated reasoning, $b = 0.24$, $t(178) = 1.97$, $p < .05$. When cognitive style was added to the regression cognitive ability no longer predicted motivated reasoning, $t(177) < 1$, $p = .84$. Instead, cognitive style became the major predictor, $b = .29$, $t(177) = 2.56$, $p = .011$. Cognitive style fully mediated the effect of cognitive ability on motivated reasoning, $b = 0.21$, $SE = 0.09$, Sobel's $Z = 2.47$, $p = .015$. An illustration of the mediation can be found in Figure 4.3.

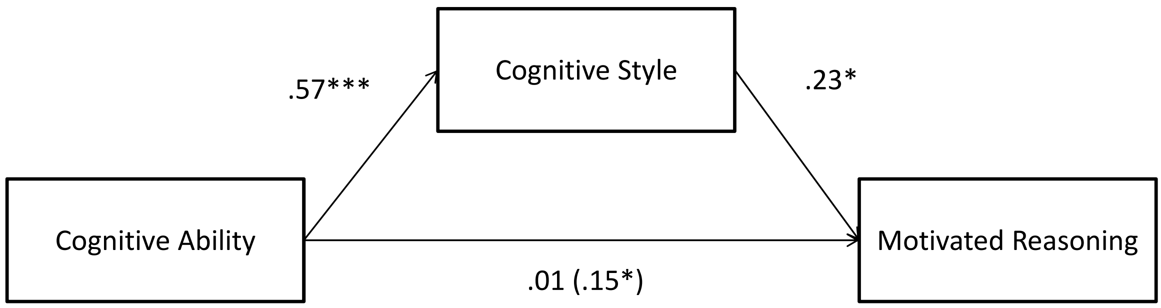


Figure 4.3. Cognitive style fully mediates the effect of cognitive ability on motivated reasoning. Path coefficients are standardised.

Response bias

We investigated whether cognitive ability mediated the link between analytic cognitive style and response bias. First, we regressed RBI on ACS. Analytic style was negatively related to response bias, $b = -1.56$, $t(178) = -4.82$, $p < .001$. When CA was added to the regression, ACS still predicted response bias, $b = -1.20$, $t(177) = -3.07$, $p = .003$, but CA did not, $b = -0.84$, $t(177) = -1.64$, $p = .10$. Consequently, cognitive ability did not mediate the effect of analytic style on response bias, $b = -0.36$, $SE = 0.22$, Sobel's $Z = -1.62$, $p = .11$.

We then investigated whether perhaps analytic cognitive style mediated the link between cognitive ability and response bias – as was the case for motivated reasoning. We started by regressing RBI on CA, showing that cognitive ability was negatively related to response bias, *b* = -1.73, *t*(178) = -3.99, *p* < .001. When ACS was added to the regression, ACS predicted response bias, *b* = -1.20, *t*(177) = -3.07, *p* = .003, but CA no longer did, *b* = -0.84, *t*(177) = -1.64, *p* = .10. It turned out that the link between CA and RBI was fully mediated by ACS, *b* = -0.89, *SE* = 0.31, Sobel's *Z* = -2.91, *p* = .004 (see Figure 4.4).

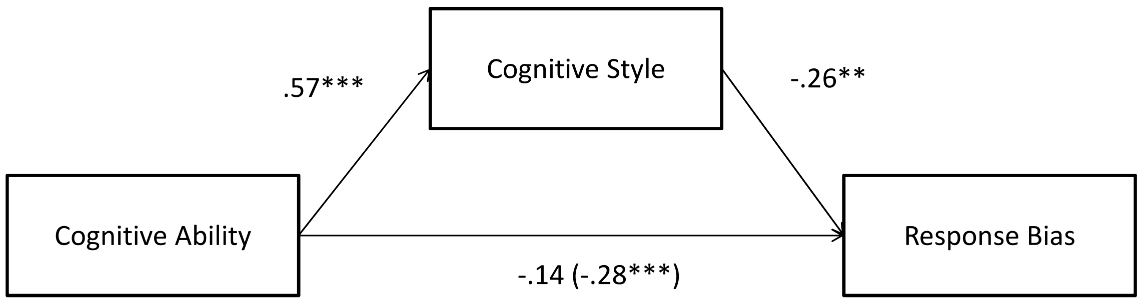


Figure 4.4. Cognitive style fully mediates the effect of cognitive ability on response bias. Path coefficients are standardised.

Path Analysis

We used path analysis to create a descriptive model of our data integrating the effects of cognitive ability and style on latency, response bias, motivated reasoning and reasoning accuracy. To take latency into account, an additional 5 outliers were removed on the basis of boxplots on response time (see Figure 4.5).

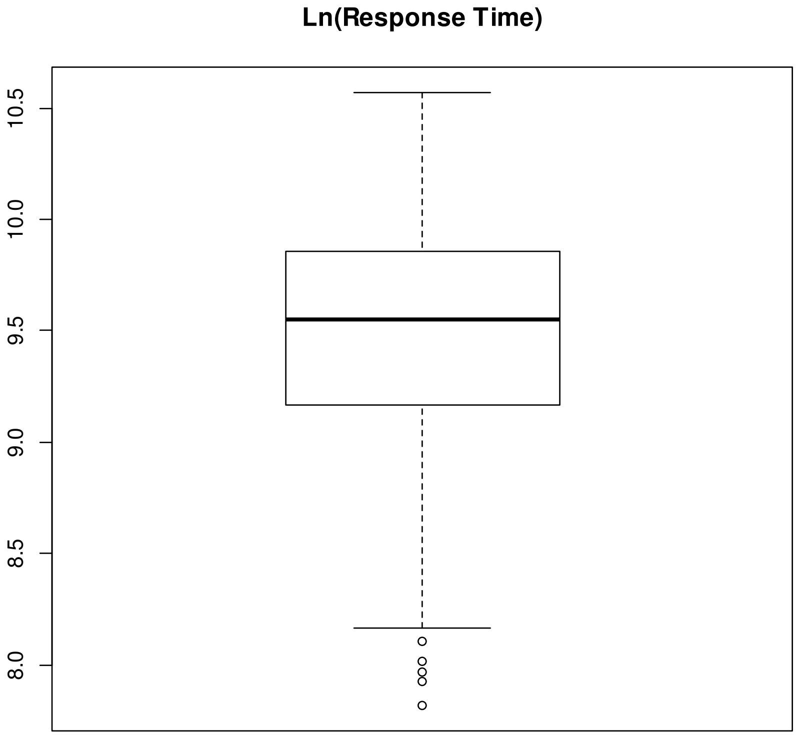


Figure 4.5. Outliers for response latency were determined using a boxplot. Including the outliers in the analyses did not change the conclusions.

In order to create our model we divided our variables into three types: cognitive ability (CA) and analytic cognitive style (ACS) were chosen as the exogenous predictor variables as they are assumed to be relatively constant due to their trait-like status. Response bias (RBI), motivated reasoning (MRI), and overall response time taken (log transformed RT, lnRT) were considered intermediate "process"-level variables, as they bear on how people reason. Finally, reasoning accuracy (A_z) was taken as the outcome variable of interest. All variables were treated as observed variables as is common in path analysis. The exogenous variables (CA and ACS) were assumed to be correlated. All endogenous variables (RBI, MRI, lnRT, A_z) were assumed to have some random measurement error associated with them. In a first step, we used maximum likelihood estimation to fit a nearly saturated model in which CA and ACS both predicted RBI, MRI, lnRT, and A_z, with lnRT predicting RBI, MRI, and A_z, and RBI and MRI predicting A_z (see Figure 4.6). The model fit the data well, χ^2 (1) = 0.63, *p* = .43.

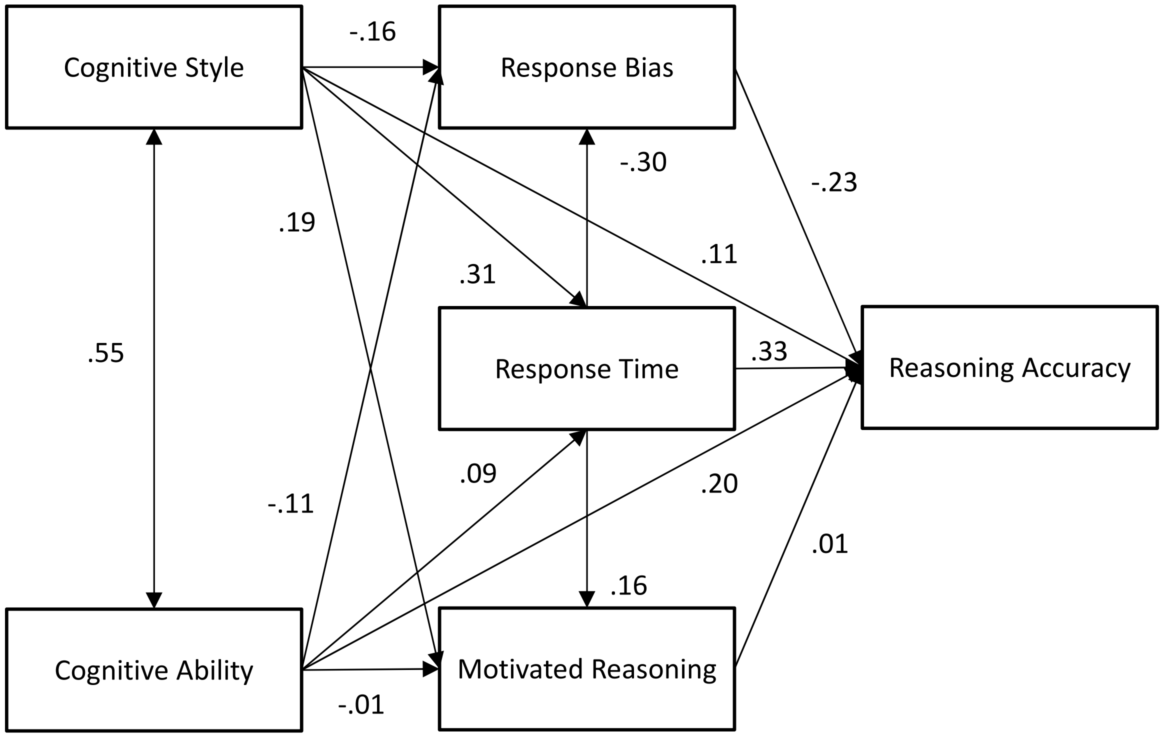


Figure 4.6. Full model relating cognitive style, ability, response bias, motivated reasoning, response time and reasoning accuracy. Path coefficients are standardised.

Inspection of the path coefficients confirmed the high correlation between style and ability. Cognitive style predicted response bias, motivated reasoning, and response time better than cognitive ability did. In contrast, cognitive ability seemed to be the better predictor of reasoning accuracy. Finally, longer response times were linked to better reasoning accuracy, as was less response bias. In an attempt to create a more parsimonious model of the data, we pruned some of the weaker paths from the full model (Figure 4.7).

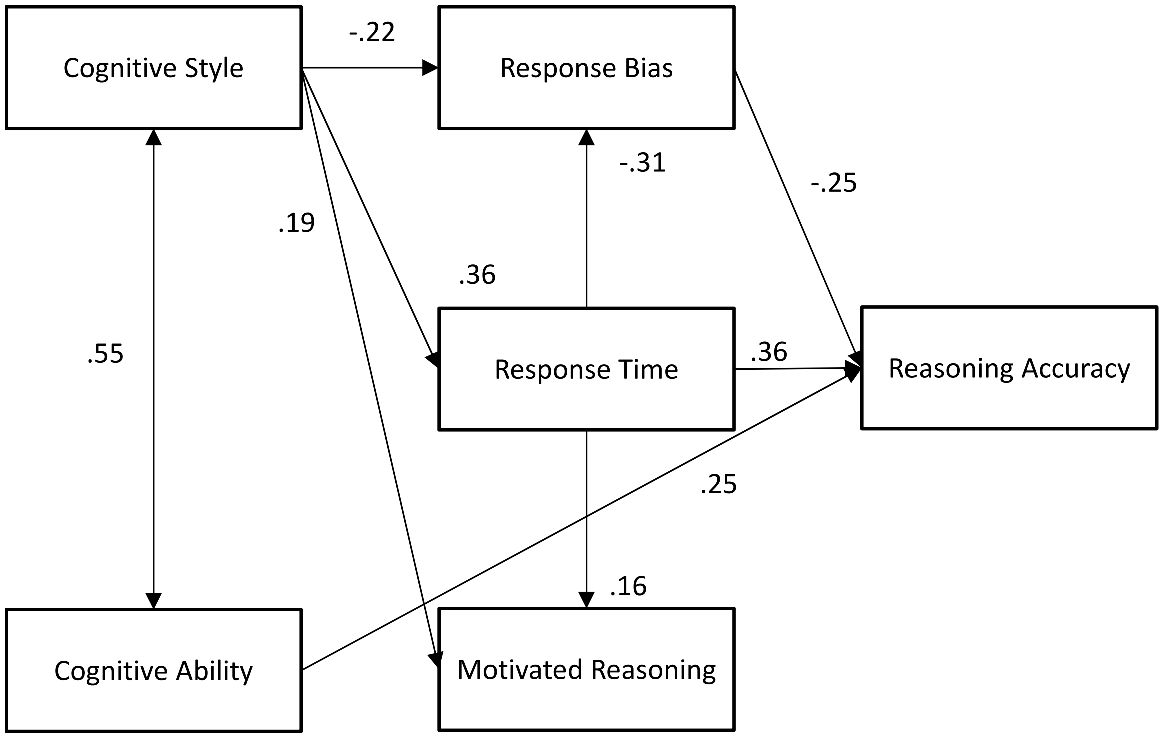


Figure 4.7. Reduced model relating cognitive style to response bias, motivated reasoning, and response time taken. Cognitive ability is linked to reasoning accuracy. Response time predicts motivated reasoning, response bias, and reasoning accuracy. Path coefficients are standardised.

To test whether these restrictions significantly reduced the fit of the model, we conducted a likelihood ratio test. Imposing the outlined restrictions did not significantly reduce the fit of the model, $\Delta\chi^2$ (5) = 5.14, p = .40. The reduced model fit the data well, $\chi^2(6)$ = 5.66, p = .45.

4.4.3 Discussion

The aim of our final experiment was to investigate whether cognitive ability or cognitive style was the more important predictor of motivated reasoning. The data demonstrated that cognitive style was a better predictor of motivated reasoning than cognitive ability. The same held true for response bias, with higher levels of analytic cognitive style predicting reduced levels of response bias.

These conclusions were reached by combining various analytic techniques. We used the ROC procedure to replicate our previous findings which were based on the forced choice method, providing more support for the idea that motivated reasoning was mainly engaged in by a subgroup of high ability and/or style participants. These analyses further revealed that response bias was fairly universal, although the effect was larger for the low style/ability participants. Next, mediational analyses demonstrated that analytic style appeared to be responsible for the link between cognitive ability and motivated reasoning, and the link between cognitive ability and response bias. Finally, path analysis was used to bring together the mediational analyses and provided the additional insight that cognitive ability was the main predictor of actual reasoning accuracy, in contrast to analytic style, which mainly influenced the "processing variables", i.e., how long participants took to respond, how response biased they were, and how much motivated reasoning they engaged in. Even though we must be cautious with this interpretation given the caveats of mediational and path analysis (Bullock, Green, & Ha, 2010), we want to stress that this model is psychologically viable. According to this individual differences model of belief bias, high ability is linked to increased reasoning accuracy, whereas high style is linked to increased motivated reasoning and decreased response bias. Adding to the psychological plausibility, the path analysis demonstrated that these effects of cognitive style were partially mediated by the amount of time taken to respond. We discuss these findings in more detail and in light of the other findings in this chapter in the general discussion.

4.5 General Discussion

In the current chapter we focused on the impact of individual differences in cognitive ability and analytic cognitive style in an attempt to test our proposed individual differences account of belief bias. In Experiment 9 we used the forced choice reasoning method to replicate Experiment 8's finding that high cognitive ability is linked to motivated reasoning. In Experiment 10, we used the forced choice method to demonstrate that cognitive style also predicted motivated reasoning. This finding was replicated using the ROC method in Experiment 11. More importantly, in Experiment 11 we demonstrated that analytic style, rather than cognitive ability, was the main predictor of motivated reasoning and response bias. These findings are incompatible with the response bias account of belief bias, according to which motivated reasoning is not a component of belief bias. To the contrary, our findings show that motivated reasoning is indeed a significant component of belief bias. To account for this pattern of results we proposed a modified individual differences account of belief bias in which analytic cognitive style predicts the tendency to engage in more motivated reasoning and to resist response bias. This account further suggests that cognitive style partially has its effects through the recruitment of more time to respond, a proxy of cognitive reflection (Frederick, 2005).

4.5.1 Individual Differences Matter

The presented experiments also provide some important methodological insights for the study of reasoning (and perhaps even cognitive psychology in general). The conclusions drawn from the experiments in this chapter would have been rather different had we not explicitly taken individual differences into account. Consider for instance the following hypothetical scenario. If we had run Experiment 9 with exactly the same participants but without taking individual differences in cognitive ability into account the very straightforward and clear conclusion would have been that participants engaged in response bias ($M_{\text{non-conflict}}$ = .74, M_{conflict} = .67), $F(1, 95) = 6.89$, $p = .010$, $\eta_p^2 = .067$, but not motivated reasoning ($M_{\text{believable}}$ = .73, $M_{\text{unbelievable}}$ = .74), $F(1, 95) = 1.39$, $p = .24$, $\eta_p^2 = .014$, providing converging evidence for the response bias account using a novel method within the SDT paradigm. Next, having established that motivated reasoning was absent in Experiment 9, perhaps we would have been interested in attempting to replicate these findings using a blocked design. Once again, the analysis of Experiment 10 would have indicated that participants engaged in response bias ($M_{\text{non-conflict}}$ = .76, M_{conflict} = .63), $F(1, 60) = 10.89$, $p = .002$, $\eta_p^2 = .15$, but not motivated reasoning ($M_{\text{believable}}$ = .71, $M_{\text{unbelievable}}$ = .73), $F(1, 60) = 1.39$, $p = .24$, $\eta_p^2 = .023$.

Finally, having firmly established strong evidence compatible with the response bias account using the forced choice paradigm, we may have been interested in doing one large final study using the ROC method. As in the previous two experiments, we would have concluded that participants engaged in response bias, ($c_{a\ \text{believable}}$ = -0.68, $c_{a\ \text{unbelievable}}$ = 0.05), $F(1, 189) = 105.85$, $p < .001$, $\eta_p^2 = .36$. In contrast with the previous two experiments, however, the analysis would have indicated that participants engaged in a very small degree of motivated reasoning ($A_{z\ \text{believable}}$ = .66, $A_{z\ \text{unbelievable}}$ = .68), $F(1, 189) = 4.31$, $p = .039$, $\eta_p^2 = .02$. Perhaps we would have concluded that the non-simultaneous forced choice method eliminates response bias sometimes (E4) and motivated reasoning at others (E9, E10). Similarly, we may have concluded that the ROC method always leads to response bias, but only occasionally to motivated reasoning (E5, E11). Furthermore, the motivated reasoning effects were so small that they are probably psychologically insignificant. The main conclusion may then have been that Dube et al. (2010) were accurate in their conclusions and that belief bias is just a response bias.

4.5.2 Latency, Confidence and Dual Process Theories

We also collected and analysed response latency and confidence ratings, although this was not the main focus of our analyses. There was some consistency, although the results were not overly indicative in providing evidence in favour or against various accounts. In general, it was found that the higher ability and/or style participants took more time to respond than the lower ability/style group. Furthermore, valid problems were processed faster than invalid problems, pointing to a potential fluency effect (Morsanyi & Handley, 2012). The most striking finding was that problem type and believability did not appear to consistently impact latency. In E9, for instance, it was found that participants took more time to respond to believable problems than to other problem types. In contrast, the analysis of E10 revealed that participants responded more quickly to non-conflict compared to believable problems. The only consistent finding was that participants did not appear to take more time to respond to the unbelievable problems. At face value, this could be considered evidence against motivated reasoning accounts of belief bias which suggest that unbelievable conclusions should result in additional analytic processing compared to believable conclusions, leading to increased reasoning accuracy (e.g., Thompson, Newstead, & Morley, 2011; Thompson, Striemer, Reikoff, Gunter, & Campbell, 2003). We failed to replicate Thompson et al.'s findings, although one explanation is that our crude measure of response time did not distinguish between reading, reasoning, and rationalising time. It is possible, for instance, that believable conclusions occasionally lead to drawn-out response times due to an increased confirmation tendency. If anything, the findings were more consistent with Stupple, Ball, Evans, & Kamal-Smith's (2011) modified selective processing account, which predicts that participants with a higher reasoning aptitude have different response time profiles than pororer reasoners. The path analysis in Experiment 11 confirmed that participants who took longer to respond were less likely to show belief bias as response bias, more likely to engage in motivated reasoning, and more likely to get the correct response. Importantly, though, response time was linked to cognitive style, not cognitive ability – even though the latter is typically considered a more potent predictor of reasoning ability (Stanovich & West, 2008).

These findings are compatible with dual process theories (DPT) of reasoning according to which cognitive ability and cognitive style are determinants of belief bias and reasoning aptitude. Unfortunately, DPT is often treated as a meta-theory which does not make many concrete predictions about the effect of the various individual differences and their interrelation (Evans & Stanovich, 2013a; 2013b). These types of predictions are usually reserved for particular instantiations of the theory (e.g., some of the traditional theories of belief bias such as selective processing theory). For instance, even though DPT typically acknowledges that style and ability are correlated with each other, as well as with response bias, the presented results in the literature did not lead to modifications of the traditional belief bias theories (Sá et al., 1999). There is also no consensus in the literature whether cognitive ability is meant to lead to increased (Newstead et al., 2004) or decreased (Quayle & Ball, 2000) motivated reasoning. In the current chapter, we have attempted resolve these inconsistencies by applying the SDT method. In doing so, we proposed a modified individual differences account of belief bias, according to which cognitive style is linked to both components of belief bias, and cognitive ability is linked to reasoning performance, with reasoning time playing a significant role in explaining the former (see Figure 4.7).

The main conclusions of this chapter are that ignoring individual differences in the study of belief bias can lead to incorrect conclusions, and that cognitive style appears to be the major predictor of the degree of motivated reasoning and response bias. We now turn to Chapter 5 in which we present a general discussion of the findings in this dissertation and their wider implications.

Chapter 5: General Discussion

5.1 Introduction

The aim of this dissertation was to further our understanding of human reasoning. We discussed the various systematic biases which impact on deductive reasoning, focusing specifically on belief bias – the interference of induction with deduction (Wilkins, 1928). Various accounts of belief bias were introduced, all of which have in common that their main aim is to provide a psychological theory for two reliable empirical findings (Evans, Barston, & Pollard, 1983). First, all theories must account for the main effect of beliefs on conclusion endorsement, or the fact that believable conclusions are accepted more than unbelievable ones (response bias). Second, all theories must explain the interaction between logic and belief, or the finding that people appear to be better at discriminating between valid and invalid syllogisms when such arguments have unbelievable conclusions (motivated reasoning).

We then introduced the response bias account of belief bias by Dube, Rotello, and Heit (2010) which explains the main effect of beliefs in terms of a response bias and the logic x belief interaction in terms of a Type 1 error. Dube et al. used SDT to demonstrate that the traditional analysis of endorsement rates typically used to investigate belief bias is flawed, because it presumes linear receiver operating characteristics (ROC) curves. In contrast to this assumption, reasoning ROCs turned out to be curvilinear (Dube et al., 2010; 2011; Heit & Rotello, 2005; 2010; Rotello & Heit, 2009). The implication of their argument was that all traditional theories of belief bias are incorrect. In response to this strong conclusion, Klauer and Kellen (2011) applied an alternative measurement model based on the multinomial processing tree (MPT) framework which could also account for the curvilinear ROCs. This model interpreted the logic x belief interaction in terms of motivated reasoning. Dube et al. (2011) rejected this alternative measurement method on the account of it being too flexible and less psychologically viable than the SDT model. In sum, the belief bias debate centred on the question of whether motivated reasoning is a component of belief bias or not.

We adopted the SDT approach in order to advance the belief bias debate. In Chapter 2, we extended the SDT approach by developing a novel forced reasoning method which allowed us to measure response bias and motivated reasoning directly. In the first three experiments two arguments were simultaneously presented, one of which was always valid and one of which was always invalid. In Experiment 1, we eliminated the possibility for participants to show response bias, leaving open the possibility for motivated reasoning only. Consistent with Dube et al.'s response bias account, motivated reasoning was not found. In Experiment 2, we reintroduced the possibility for participants to show response bias, in an attempt to find converging evidence for the response bias account. Contrary to our predictions, response bias was not found. We hypothesised that perhaps the forced choice method somehow altered the participants' reasoning strategy, effectively eliminating response bias (and perhaps motivated reasoning).

We proposed two ways in which the simultaneous presentation may have altered the reasoning process. According to the belief suppression hypothesis, presenting the response bias problems randomly intermixed with the motivated reasoning problems caused the participants to realise that beliefs and validity are unrelated, leading them to actively suppress their prior beliefs and to using an alternative reasoning process. In Experiment 3, this belief suppression hypothesis was disconfirmed, with no response bias occurring even though the response bias problem types were presented in the absence of the motivated reasoning problems. According to the structural focus hypothesis, the simultaneous problem presentation cued premise to conclusion reasoning instead of the default conclusion to premise reasoning strategy (Morley, Evans, & Handley, 2004), causing both components of belief bias to be eliminated. In Experiment 4, we tested the structural focus hypothesis: non-simultaneous problem presentation was used to induce conclusion to premise reasoning, and evidence consistent with motivated reasoning was found. This finding provided evidence incompatible with the response bias account of belief bias.

In Chapter 3 we aimed to resolve the inconsistency between Experiment 4 – in which evidence consistent with motivated reasoning was found – and Dube et al.'s key experiment in which motivated reasoning was not found. Although both experiments explicitly employed methods situated within the SDT framework, one possible explanation for the discrepancy was a difference in the specific method (forced choice v. ROC). Throughout Chapter 3 we used the ROC method to eliminate this possibility. In Experiment 5, we investigated whether syllogism complexity was a determinant of motivated reasoning. Motivated reasoning was found for complex (multiple model), but not simple (one model) syllogisms, but response bias was found for both problem types. We hypothesised that Dube et al.'s participants may have differed from ours a certain way (e.g., in motivation or its proxy: time taken to respond). In Experiment 6, we investigated the impact of a response time manipulation. Inconsistent with previous research (Evans & Curtis-Holmes, 2005), the speeded task manipulation had no effect on either component of belief bias. Even more confusing was the finding that motivated reasoning did not occur in the self-paced condition, replicating Dube et al. (2010). In Experiment 7 we investigated the effect of pragmatic instructions on response bias and motivated reasoning, hypothesising that such instructions might lead to more response bias, and potentially more motivated reasoning. Instructions did not impact on either component of belief bias. Furthermore, we failed to replicate Experiments 4 and 5, finding once again that motivated reasoning was absent. An exploratory analysis of Experiment 7 suggested that only those participants who took more time to reason engaged in motivated reasoning.

Inspired by the idea that different participants may have been reasoning in different ways, in Experiment 8 we investigated the effect of individual differences in cognitive ability alongside a response time manipulation. Motivated reasoning was engaged in mainly by those of higher cognitive ability, whereas response bias was universally present. The response time manipulation successfully eliminated (or even reversed) motivated reasoning for the higher ability subgroup, in contrast to Experiment 6 where the manipulation had no effect. In light of these findings we proposed an individual differences account of belief bias, according to which everybody shows response bias, but only those of higher cognitive ability engage in motivated reasoning. This account provided a potential explanation for the inconsistent findings in the previous experiments: perhaps motivated reasoning was present in certain experiments (Experiments 4 and 5) but absent in others (Experiments 6 and 7) because a smaller proportion of higher ability participants was sampled in the latter compared to the former.

In Chapter 4 we further tested the individual differences account of belief bias, employing both the forced choice and the ROC method. In Experiment 9 we successfully replicated the finding that motivated reasoning was engaged in mainly by those of higher cognitive ability using the non-simultaneous forced choice reasoning method, demonstrating that the effect was robust to variations in the (SDT) method. In Experiment 10, we demonstrated that cognitive style as measured by the cognitive reflection test (CRT; Frederick, 2005) also predicted motivated reasoning. In Experiment 11, we used a meditational analysis in combination with the ROC method to demonstrate that the effect of cognitive ability on motivated reasoning and response bias was mediated by cognitive style. Finally, path analysis was used to present a statistically and psychologically viable individual differences model of belief bias. This model is discussed in more detail in the next section.

5.2 Implications

As outlined in the summary above, the main aim of this thesis was to further the belief bias debate which centred on the question of whether motivated reasoning is a component of belief bias, or whether a pure response bias is sufficient to explain the data. We now turn to a discussion of the wider methodological, empirical, and theoretical implications of this work.

5.2.1 Methodological

The use of SDT as a more appropriate way of measuring human reasoning played a central role in the experiments in this dissertation. This approach was advocated by Dube et al. (2010), who demonstrated that the relationship between hits and false alarms in reasoning is curvilinear and that endorsement rate analysis is inappropriate, suggesting that the use of a more appropriate measurement method such as SDT led to increased consistency and parsimony, showing that belief bias was just a response bias. Using a more appropriate measurement method was an important step forward, but some issues still remained. For instance, using the appropriate measurement method did not guarantee more consistent measurements: even though SDT (Dube et al., 2010) and MPT (Klauer & Kellen, 2011) were both able to accommodate curvilinear ROCs, disagreement on the status of motivated reasoning remained depending on the chosen method.

We actively avoided getting into the SDT – MPT debate (e.g., Province & Rouder, 2012), but we had some reasons to prefer SDT in this dissertation. First, SDT is a very general theory of human decision making which has been successfully applied in other important domains of cognitive psychology, most notably memory and psychophysics (Pazzaglia, Dube, & Rotello, in press). Consequently, SDT offers a general framework which is applicable in any task where a marked preference for one item class over another (i.e., a response bias) may be confounded with sensitivity or accuracy (Green & Swets, 1966; Macmillan & Creelman, 2005). The case of belief bias lends itself very well to this approach, given the marked preference of believable over unbelievable syllogisms (e.g., Evans et al., 1983). In contrast, the MPT method requires a novel instantiation of the model for each problem, with the only common ground being that discrete detect-states can be reached with a certain probability of a threshold being exceeded (Luce, 1963). Second, the SDT model's ability to relate confidence ratings to accuracy in order to enrichen the data via the mechanism of criterion placement is psychologically justified on the basis of various experiments on metacognition (Thompson, Prowse Turner, & Pennycook, 2011). Results from such studies suggest that confidence or the so-called feeling of rightness plays an important role in reasoning and judgement, for instance in deciding whether or not to stick with the initial response. In contrast, for MPT to accommodate for curvilinear ROCs, the (arguably less psychologically plausible) assumption needs to be made that participants differ in the way in which they use the confidence rating scale (Klauer & Kellen, 2011). Finally, model recovery simulations demonstrated that the MPT model was more flexible than the SDT model, given that the former was better at fitting data which was generated from the latter than vice versa (Dube, Rotello, & Heit, 2011).

We extended the SDT method using the forced choice reasoning procedure to avoid the pitfalls of traditional modeling – as well as endorsement rate analysis. The major contribution of the forced choice method was its ability to provide methodologically correct estimates of accuracy (Macmillen & Creelman, 2005). The method allowed us to measure response bias and motivated reasoning directly by comparing accuracy between different problem types. Consequently, the various theories of belief bias – including the response bias account – could be tested. Using a better measurement method (SDT) which avoided the pitfalls of modeling (forced choice) turned out not to be sufficient to decrease the inconsistency, however: motivated reasoning and response bias appeared and disappeared between experiments, even though the conditions were identical. Only when we took individual differences into account a major leap forward in consistency was found: higher ability participants were more likely to engage in motivated reasoning than lower ability participants, regardless of the specific method used (i.e., ROC or forced choice). Consequently, we argue that the main methodological implication of this dissertation is that individual differences in cognitive ability and cognitive style need to be taken into account when investigating belief bias. Merely using a more appropriate measurement method alone was not sufficient to resolve the question whether motivated reasoning is a component of belief bias or not, although neither was accounting for individual differences when using an inappropriate measurement method (see in this respect the inconsistent observation of the logic x belief interaction as a function of cognitive ability found in Quayle & Ball, 2000 and Newstead et al., 2004).

The key to advancing our understanding of belief bias – and reasoning in general – turned out to be a combination of using the appropriate measurement method and accounting for individual differences. Future research on human reasoning would also benefit from taking these methodological considerations into account.

5.2.2 Empirical

The main empirical question of this dissertation was whether motivated reasoning could be observed in syllogistic reasoning if the appropriate measurement method was used. The answer turned out to be yes, but only under certain conditions: motivated reasoning was observed only for people reasoning about sufficiently complex arguments, who were provided with adequate time to respond, under a certain set of standard instructions, and who possessed above average levels of analytic cognitive style and/or ability. Even though this was the main empirical finding, some additional interesting observations were made.

The first three experiments presented in this dissertation employed the simultaneous forced choice reasoning method. In this method, two syllogisms were presented side by side and the participants were instructed to choose the valid one. Even though the simultaneous forced choice method did not actively increase our understanding of belief bias in terms of which account is more viable, it had the unexpected side-effect of completely eliminating both components of belief bias, while still retaining above chance reasoning performance. The finding that forced choice reasoning can eliminate both components of belief bias is novel, and particularly interesting given the inability of instructional manipulations to fully eliminate belief bias (Evans, Newstead, Allen, & Pollard, 1994; Newstead, Pollard, Evans, & Allen, 1992). The simultaneous forced choice method may provide to be useful in future research aiming to educate people on logical reasoning by demonstrating that logic and beliefs are unrelated.

The remaining forced choice experiments, which employed a non-simultaneous problem presentation, demonstrated that the forced choice method is capable of directly measuring response bias and motivated reasoning whilst conforming to the underlying assumptions of ROC curvilinearity. It turned out that response bias was engaged in by all participants, although follow-up tests suggested that the effect was particularly strong for the lower style/ability participants. Nevertheless, these findings replicated and confirmed the classic finding that belief bias as response bias is pervasive and universal (Wilkins, 1928).

Syllogism complexity has been used as a means of distinguishing between the various theories of belief bias. For instance, according to Mental Models Theory (Oakhill et al., 1989) and Selective Processing Theory (Evans et al., 2001; Klauer et al., 2000), the logic x belief interaction should be absent for simple (i.e., one-model) syllogisms. In contrast, Selective Scrutiny predicts that the logic x belief interaction should be present for all problems regardless of their complexity (Evans et al., 1983). The data with respect to the logic x belief interaction in one-model syllogisms is inconsistent, as it was absent in certain experiments (e.g., Newstead et al., 1992; Klauer et al., 2000) and present in others (Evans & Pollard, 1990; Glinsky & Judd, 1994). In Experiment 5, we demonstrated that motivated reasoning was absent for simple problems when SDT was used. Although one potential issue with this finding is that individual differences were not taken into account, the fact that motivated reasoning occurred for the complex condition suggests that the sample was of above average cognitive ability and/or style. Nevertheless, additional research specifically taking these variables into account is required to verify the finding.

Like complexity, response time is frequently manipulated in reasoning to test various theories (e.g., Evans, Handley, & Bacon, 2009; Schroyens, Schaeken, & Handley, 2003). For instance, according to Misinterpreted Necessity (Evans et al., 1983), the logic x belief interaction originates from the application of a belief heuristic for the very difficult indeterminately invalid problems, and limiting response times should increase the use of such heuristics, resulting in a larger interaction for the speeded group. In contrast, MMT and SPT predict that limiting the response time available should reduce the ability to engage in motivated reasoning, resulting in an absence of the interaction. Evans and Curtis-Holmes (2005) found that motivated reasoning was eliminated, consistent with predictions from MMT and SPT. The results also indicated that belief bias as response bias increased in magnitude under a response time limit. This latter finding was taken in support of dual process theories (DPT) of reasoning, as it suggests that the influence of beliefs originates from a heuristic process which takes time to be overridden by additional processing. One caveat of this study was that the results were based on a traditional analysis of endorsement rates. We investigated the effect of the response time manipulation in combination with the more appropriate ROC method (Experiment 6) and taking individual differences into account (Experiment 8). Contrary to Evans and Curtis-Holmes' findings, limiting the amount of time available to respond did not increase the magnitude of the response bias. In fact, the only effect of the response time limit on criterion placement appeared to be that the high ability group was slightly more conservative in the speeded condition, regardless of beliefs. These findings suggest that the effects of response time on belief bias might not be as well understood as previously thought, with the time limit mainly having an effect on the motivated reasoning component of belief bias, and only for the higher ability/style participants.

A similar story can be told about the effect of pragmatic instructions on belief bias. In Experiment 7, we investigated the impact of an instructional manipulation on response bias and motivated reasoning. Previous research demonstrated that pragmatic instructions led to increased response bias in a causal conditional reasoning task, regardless of ability (Evans, Handley, Neilens, & Over, 2010). The conclusions from these studies were based on analysis of endorsement rates. In contrast, using the SDT method, we did not find any effect of instructions. The cause for this discrepancy can be two-folded. Either the difference in the measurement method was responsible, or the fact that we did not look at the impact of individual differences obfuscated the results. Evans et al. (2010)'s results suggest that, perhaps lower and higher ability/style groups make use of pragmatic instructions in qualitatively different ways. Future research using a more appropriate measurement method and taking individual differences into account will indicate whether this is the case. Regardless what such research might find, the key empirical findings of Experiments 6 (response time) and 7 (instructions) is that they underscore the pervasiveness of belief bias as response bias.

Like in much of the previous research on belief bias (e.g., Evans et al., 1983), the bulk of our theorising focused on the effect of conclusion believability on reasoning accuracy (motivated reasoning) and conclusion endorsement (response bias). However, more recently, researchers interested in reasoning and DPT have started focusing more on processing variables such as response time taken and confidence ratings. This research has suggested that decreased confidence and increased response times are indicative of the implicit detection of a conflict between the normative (e.g., responses in line with logic or probability theory) and the heuristic (e.g., responses in line with prior beliefs) task characteristics (e.g., De Neys & Glumicic, 2008; Pennycook, Cheyne, Koehler, & Fugelsang, 2013; Thompson et al., 2003; Thompson et al., 2011). Even though it was not our primary focus, we also collected response latencies and confidence ratings. In fact, the latter were an essential part of the ROC method. In contrast to this promising trend advocated by the studies above, our response time data did not provide evidence consistent or inconsistent with the response bias account: beliefs did not impact on response time taken, at least when the ROC method was used (Experiments 5, 6, 7, 8, 11).

Response time did occasionally differ as a function of believability in the forced choice experiments, but the findings were inconsistent. Compared to the other problem types, unbelievable problems were responded to more quickly sometimes (Experiment 1), but equally fast at others (Experiments 2-4). Believable problems were responded to more slowly in one study (Experiment 9), but not in other ones (Experiments 1 – 4, Experiment 10). Neutral contents elicited longer response times in one study (Experiment 3), but not in a different one (Experiment 1). The only interesting consistent response time finding was that high style/ability participants took longer to respond than their counterparts (Experiments 8 – 11), possibly due to their increased tendency towards cognitive reflection. Furthermore, this additional time significantly increased reasoning accuracy and motivated reasoning, but decreased response bias (Experiment 11). Our findings suggest that response time is a less optimal dependent variable for belief bias theorising, but that it can play a key mediational role. One potential reason why its usefulness as a dependent variable is reduced might be found in the fact that response time is not identical to reasoning time, due to the fact that reasoning generally takes a lot of time and effort. Response times in reasoning are probably a fairly crude measure given that it is difficult to disentangle the time spent reading, reasoning, and rationalising. More fine-grained measures of response time which somehow manage to differentiate between these response phases, possibly in combination with eye-tracking data (Ball, Phillips, Wade & Quale, 2006) may turn out to fare better.

Similar inconsistencies were found for the confidence rating data in the forced choice experiments. In the non-simultaneous forced choice experiments, believable problems elicited lower confidence in one case (Experiment 4), higher confidence in another one

(Experiment 10), or had no effect whatsoever (Experiment 9). High ability participants were more confident than low ability participants (Experiment 9), mimicking Shynkaruk and Thompson's (2006) findings, but no confidence differences as a function of cognitive style were found (Experiment 10). Slightly more consistency was found in the simultaneous forced choice experiments, with neutral problems eliciting lower confidence than believable and unbelievable (Experiment 1) or non-conflict and conflict (Experiment 3) problem types. No confidence differences between the four problem types were found (Experiment 2). Taken together, these findings suggest that confidence ratings on their own did not contribute many interesting insights into our understanding of belief bias. The main function of confidence ratings appears to elevate the usefulness of endorsement rates allowing us to estimate SDT parameters using the ROC procedure.

5.2.3 Theoretical

The starting point of this dissertation was a theoretical debate about the status of motivated reasoning. According to the response bias account, belief bias is just a response bias and motivated reasoning is not a significant component of belief bias. We demonstrated that motivated reasoning is a part of belief bias in syllogistic reasoning. Consequently, the main theoretical contribution is that the response bias account of belief bias proposed by Dube et al. (2010) is insufficient to account for the current pattern of results.

A more constructive implication for the belief bias theories is that cognitive ability and cognitive style need to be taken into account. We have provided a jumping-off point by proposing the individual differences account of belief bias (Figure 5.1). This account consists of a psychologically and statistically viable pathway model which proposes a relationship between cognitive ability, cognitive style, response time taken, motivated reasoning, response bias, and reasoning accuracy. This model suggests that cognitive ability is the better predictor of actual reasoning performance, with analytic cognitive style being the better predictor of motivated reasoning and response bias. Furthermore, this model suggests that style partially has its effect through the recruitment of additional cognitive reflection time, via which it also leads to additional reasoning accuracy. Importantly, the model specifically acknowledges the fact that analytic style and cognitive ability are highly correlated.

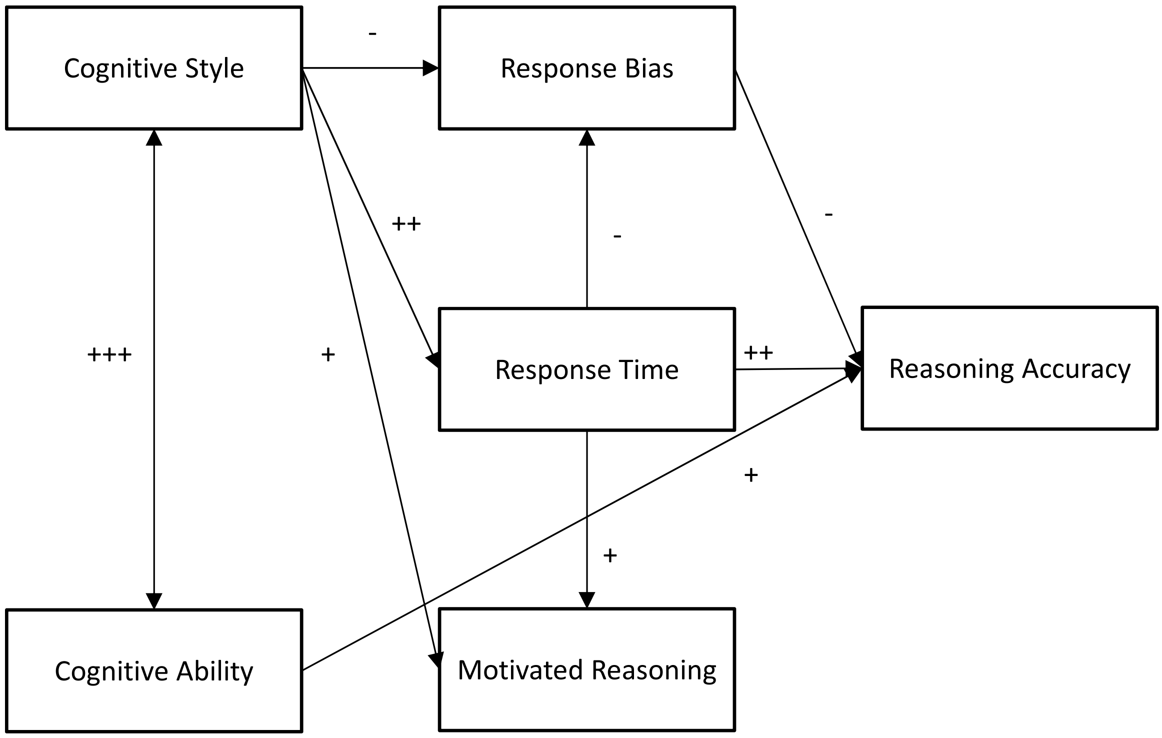


Figure 5.1. Individual differences model of belief bias.

Rather than providing an actual detailed process-level theory of belief bias such as Mental Models Theory (MMT; Oakhill et al., 1989) or Selective Processing Theory (SPT; Evans et al., 2001), the individual differences model of belief bias can be seen more as an overarching framework which can unite the more detailed theories of belief bias, none of which specifically account for individual differences in style and ability. For example, it is possible that lower style participants operate in a pure response bias manner as proposed by Dube et al.'s (2010) response bias account, engaging only in limited actual reasoning, possibly relying instead on certain heuristics such as those proposed by the Probability Heuristics Model (PHM; Chater & Oaksford, 1999; Oaksford & Chater, 2007; 2009) to achieve a modicum of logical reasoning competence. Participants of higher analytic cognitive style and sufficient cognitive ability, in contrast, might operate in a manner consistent with SPT (adopting a negative testing strategy for unbelievable problems and a positive testing strategy for believable problems) or MMT (searching for additional counterexamples in the face of an unbelievable conclusion). The model can also explain how certain very high ability participants respond in a perfectly normative manner by performing an exhaustive search for all the mental models regardless of believability, or by abstracting the problem and applying the relevant logical rules.

The main theoretical implication of Dube et al.'s (2010; 2011) work was that all the traditional theories of belief bias were incorrect, because these theories provided a psychological explanation (i.e., motivated reasoning) for a statistical artefact (the logic x belief interaction, a type 1 error). Given that our work has established that motivated reasoning is indeed a component of belief bias, we now revisit the original theories of belief bias which were introduced earlier and evaluated how they fare in light of the current findings.

Traditional Belief Bias Theories

According to Selective Scrutiny (SS; Evans et al., 1983), participants accept believable conclusions but engage in additional reasoning for unbelievable conclusions. As such, the account predicts a universal response bias and motivated reasoning affecting all types of syllogisms, including the simplest ones. Evidence incompatible with SS was found in Experiment 5, in which we demonstrated that motivated reasoning was eliminated for simple (one model) syllogisms. More generally, it was also found that motivated reasoning did not occur in Experiments 1 and 3 which used the forced choice method. SS does not specify that the belief bias effects originate from conclusion to premise reasoning. Consequently, there is no reason why simultaneously presenting two unbelievable arguments should lead to the elimination of motivated reasoning. SS has been rejected before in previous studies (e.g., Klauer, Musch, & Naumer, 2000), but these experiments did not account for ROC curvilinearity. Our current findings alongside the earlier findings cement the idea that SS has failed as a model of belief bias.

Misinterpreted Necessity (MN; Evans et al., 1983) asserts that participants do not grasp the concept of necessity. Consequently, when presented with complex syllogisms in which multiple models of the premises need to be constructed to reach the correct conclusion (e.g., indeterminately invalid – aka multiple model – syllogisms), participants instead respond on the basis of believability, resulting in motivated reasoning due to the increased rejection of unbelievable invalid problems. Note that the term motivated reasoning is a bit of a misnomer in this case: the observed accuracy effect is actually driven by the application of a belief heuristic (participants are giving the right response for the wrong reasons). One prediction from this account is that imposing a response time limit should lead to *more* motivated reasoning, given that the effect is thought to originate from the use of a belief heuristic, and that more heuristic responding supposedly occurs when participants are placed under a time limit (Evans & Curtis-Holmes, 2005). We demonstrated that this does not happen (Experiment 6), or even that the opposite happens (Experiment 8).

Metacognitive Uncertainty (MU; Quayle & Ball, 2000) is an extension of MN which suggests that motivated reasoning occurs because the working memory resources of the average participant are exceeded by the requirements of indeterminately invalid syllogisms. According to this theory, then, motivated reasoning should be absent for those of higher working memory capacity, a correlate of cognitive ability (Conway, Kane, & Engle, 2003). Experiments 8, 9, and 11 demonstrate that the reverse occurs: participants of higher cognitive ability (and thus presumably higher WMC) show more motivated reasoning.

The Modified Verbal Reasoning Theory (MVRT; Thompson et al., 2003) originated from the finding that people spend more time reasoning about believable than unbelievable conclusions. According to MVRT, participants spend more time on believable problems because they reason until a self-imposed deadline has elapsed. This deadline is thought to be longer for believable than for unbelievable problems, because the former ones are more palatable. Beliefs affect accuracy, then, because participants will base their judgement on conclusion believability if this deadline elapses and no response has been found. Consequently, for indeterminately invalid unbelievable problems, more "invalid" responses are given (due to the lower response deadline which cues a correct belief-based rejection), leading to increased accuracy. We measured response time in all of our experiments, but we did not replicate the finding that participants reasoned longer about believable compared to unbelievable problems. Consequently, the theory invokes an additional mechanism to explain an empirical result that turned out be less reliable than previously assumed. Furthermore, MVRT suggests that motivated reasoning should occur irrespective of cognitive ability or cognitive style, something that was not observed. In addition to this, one might predict that instructional manipulations or explicit response time manipulations should result in a different pattern of results, as this manipulation interferes with the self-imposed response time deadlines. Experiments 6 and 7 both provided evidence inconsistent with this prediction. Whereas future research might show that the process underlying syllogistic reasoning is verbal in nature, MVRT in its current form is not supported by the current data.

The Mental Models Theory (MMT; Oakhill, Johnson-Laird & Garnham, 1989) of belief bias explains motivated reasoning by suggesting that in the crucial stage of the evaluation of the initial conclusion, additional models are sought if the conclusion is unbelievable, leading to increased correct rejections (i.e., accuracy), but only for multiple model (that is, complex) syllogisms. This leads to two predictions. First, one model (simple) syllogisms should not lead to motivated reasoning. This exact pattern was observed in Experiment 5. Second, people of higher cognitive ability should be better at constructing the alternative models due to their increased WMC. Once again, this was confirmed by our findings in

Experiments 8, 9, and 11. One potential issue for MMT is the finding that, rather than cognitive ability, cognitive style turned out to be the major predictor of motivated reasoning. Furthermore, if it is the case that motivated reasoning results in additional time spent on model construction, then we should find evidence of drawn out reasoning times for unbelievable compared to believable problems. As mentioned above, this was not observed. MMT explains response bias by proposing that after the reasoning process is complete, a conclusion "filter" occasionally rejects unbelievable conclusions or accepts believable conclusions, for one-model problems, irrespective of cognitive ability or cognitive style. Contrary to this assumption, we demonstrated that under some circumstances response bias was severely reduced or even eliminated in a higher cognitive ability (Experiment 9) or higher cognitive style (Experiment 10) subgroup. More generally, analytic style was negatively related to response bias (E11). Despite these shortcomings, MMT provides a better explanation of the data compared to the other traditional belief bias accounts discussed above.

Selective Processing Theory (SPT; Evans et al., 2001; Klauer et al., 2000) proposes that participants reason from the conclusion to the premises (Morley, Evans, & Handley, 2004) and that they only ever construct a single mental model of the premises (Evans, Handley, Harper, & Johnson-Laird, 1999). Motivated reasoning is explained by a belief-driven alteration to the reasoning process, with unbelievable conclusions leading to a disconfirmatory strategy in contrast to believable conclusions, for which a default, confirmatory strategy is used. The negative testing strategy is thought to facilitate construction of the incompatible model, increasing the correct rejection rate of indeterminately invalid syllogisms. Believable conclusions, on the other hand, are thought to result in the incorrect acceptance of indeterminately invalid syllogisms. For valid problems, these belief based strategies have no impact – because a disconfirming model can never be found – with beliefs only occasionally making their mark in terms of a general response bias. Like MMT, SPT predicts that motivated reasoning should not occur for simple syllogisms, given that only a single model of the premises exists. Also, one prediction might be that participants of higher cognitive ability are more able to engage in the effortful (Deutsch, Gawronski, & Strack, 2006) negation-based disconfirmatory reasoning strategy cued by unbelievable conclusions. The finding that limiting response time occasionally reversed the motivated reasoning effect (e.g. Experiments 6 and 7) can also be explained by SPT: this finding suggests that those who would typically engage in disconfirmation still attempt to do so in the face of a response time deadline, but simply fail due to the increased effort required.

A modified version of SPT (Stupple, Ball, Evans, & Kamal-Smith, 2011) specifically allows for individual differences to play a role in explaining response time patterns by suggesting that qualitatively different types of reasoners exist, something which resonates well with the current findings (E7, E8, E9, E10, E11). In that sense, our individual differences account of belief bias and the findings on which it was based bears the most resemblance to this modified SPT, with the added benefit of providing an explicit relational configuration of the various individual differences which have an impact on both components of belief bias as well as reasoning accuracy. We note, however, that our model does not require that participants reason in the way proposed by SPT. As outlined above, it is equally possible that a completely different reasoning process underlies the performance pattern captured by our model. One can imagine for instance that other mechanisms explain the findings equally well. Examples of such mechanisms are the search for counterexamples proposed by MMT, the semantic encoding and re-encoding of the premises proposed by MVRT, or the use probabilistic heuristics proposed by the PHM.

The PHM (Chater & Oaksford, 1999; Oaksford & Chater, 2007; 2009) differs from the belief bias theories outlined above in various ways. The two key differences lie on the computational level (i.e., what is the normatively correct standard to compare performance to?) and the algorithmic level (i.e., what are people actually doing?). On the computational level, the theory posits that people treat the syllogistic reasoning task as a probabilistic- rather than logical reasoning task. This means that participants strive to reduce uncertainty, rather than determine necessity, accepting or drawing only conclusions which are less uninformative than the premises. On the algorithmic level, rather than using mental models or logical rules to reason, PHM proposes five fast and frugal heuristics which are used to determine the informativeness (i.e., probabilistic or p-validity) of the conclusion. The appropriateness of these heuristics is justified by probability calculus. Three heuristics are used to generate a conclusion given a set of premises: G1) the min-heuristic, G2) p-entailment, and G3) attachment. Two heuristics are used to test whether a certain conclusion is likely to be p-valid: T1) the max-heuristic and T2) the some…not heuristic (these heuristics are discussed in more detail in Chapter 1).

Our results raise the question how PHM can account for the finding that under various conditions, certain high ability participants are better at determining whether or not a conclusion is (p-) valid. One possible explanation is that certain participants under certain conditions are more likely to apply the heuristics which lead to the correct response. More specifically, the key might lie in whether the participants simply use the min-heuristic, whether they go on to use p-entailment after that, and whether they then follow this up by applying the attachment heuristic. As our experiments show, motivated reasoning occurred only for complex (multiple model) syllogisms which all have conclusions using the “some … not” quantifier. Importantly, most of these complex syllogisms featured premises with a “no” and a “some” quantifier. Application of the min-heuristic for syllogisms with “no” and “some” premises suggests that the correct conclusion should employ the “no” quantifier (because “no X are Y” is less informative than “some X are Y”, and the min heuristic proposes that the least informative premise quantifier is the preferred quantifier for the conclusion). If the participant stops after applying the min-heuristic, this will lead to correct rejections approximately half the time and incorrect rejections the other half (considering that half the problems are valid and half are invalid).

It is possible that unbelievable conclusions motivate participants to go a step further and to use p-entailment after the min-heuristic to determine that “some…not” is the next preferred conclusion quantifier (because “no X are Y” probabilistically entails that “some X are not Y”). Finally, only those with sufficiently high levels of analytic cognitive style and/or ability will proceed to apply the attachment heuristic to determine whether the order of the presented conclusion conforms to the order suggested by the attachment heuristic, leading to better performance.

An alternative way to reconcile PHM with our motivated reasoning findings lies in the test phase which follows conclusion generation. Chater and Oaksford (1999) explicitly acknowledge that the generation heuristics (G1-3) are the most important part of the model, with the test heuristics (T1 and T2) being of lesser importance. In fact, it is mentioned that the test phase might very well involve *“…the kinds of processes discussed by other theories that can account for logical performance, such as mental models […] [or] mental logic […]…”* (Chater & Oaksford, 1999, pp. 196) and that these are *“…complex processes likely to be subject to large individual variation.”* (Chater & Oaksford, 1999, p. 207). As such, it is fully possible that the bulk of the reasoning process takes the form of applying probability heuristics, with only those of higher analytic cognitive style and/or ability going beyond the default by applying the testing processes proposed by MMT or SPT. We now propose some more general considerations for future research based on these implications.

5.3 Future Research

In our previous section we have outlined the various methodological, empirical, and theoretical implications of the experiments conducted in this dissertation. We now provide a brief overview of several research topics which follow naturally from our findings.

5.3.1 PHM and Motivated Reasoning

We proposed two ways in which our individual differences model of belief bias might be reconciled with the PHM. A first possibility is that conclusion believability and individual differences impact upon the amount and the sequence of heuristics which are typically applied, which in turn leads to higher reasoning performance. Alternatively, it is also possible that high ability participants simply act more according to the principles outlined by MMT or SPT in their test phase. One way of distinguishing between these explanations could lie in the PHM's ability to account for reasoning with two additional quantifiers, namely "most" and "few". If we replicate the finding that high style/ability participants engage in motivated reasoning for all syllogisms, including those which contain the quantifiers "most" and "few", then this would suggest that the motivated reasoning effect is somehow interwoven with the application of the probability heuristics, because according to traditional logic, conclusions with statistical quantifiers are logically invalid. Consequently, if it turns out that the motivated reasoning effects generalise to the few and most quantifiers, then this suggests that the effect cannot be explained by tacking on the use of mental models or verbal reasoning in the test phase of PHM.

5.3.2 Eliminating Belief Bias

One novel finding was that the simultaneous forced choice method resulted in the elimination of both components of belief bias. This finding is particularly interesting in light of the fact that no methods are truly successful at completely eliminating belief bias. Rather than abandoning the simultaneous forced choice method in favour of its non-simultaneous counterpart, we propose three broad areas of research which could readily benefit from the application of the simultaneous method. First, from an educational point of view, the method might prove useful in teaching people about how beliefs may interfere with reasoning and decision making. Having participants solve a standard belief bias task prior to solving a simultaneous forced choice task might provide them with a visceral demonstration of how beliefs and logic are two independent concepts. Second, from a methodological point of view, the simultaneous method might prove useful for the investigation of syllogistic reasoning with meaningful contents but without any interference of beliefs. For instance, the impact of various manipulations such as complexity, response time taken, instructions, and cognitive load on syllogistic reasoning performance in the absence of belief bias could be tested, without resorting to abstract contents. Finally, the results raise the question of whether it is possible to use the simultaneous forced choice method as a prime to eliminate (one or both components of) belief bias in a consequent reasoning task. If this turns out to be the case, a next logical step would be to investigate whether the results are temporary, or whether they translate to later test phases and even different tasks.

5.3.3 Dual Process Theory

The dual process theory (DPT; e.g., Evans, 2008; Kahneman, 2011) of reasoning proposes that two types of processes exist. Type 1 processes are quick and effortlessly executed. In contrast, type 2 processes are slow and require effort (in the form of working memory resources). Belief bias has often been taken as compelling evidence in favour of DPT (e.g., Evans & Stanovich, 2013a; 2013b): the finding that participants are apparently unable to resist the influence of their prior beliefs when judging the logical validity of reasoning problems is explained in terms of a type 1 – type 2 conflict. Type 1 processes cue a quick and effortless default response on the basis of conclusion believability, which may then be inhibited using effortful type 2 processing. If such inhibition is successful, and the relevant conditions are in place (e.g., the right mindware, correct instructions, sufficient time, adequate levels of cognitive ability and style), then the participant may proceed by computing the normatively correct response by engaging in logical reasoning – which also requires type 2 processing. A strong point of DPT is that various predictions about the impact of certain manipulations can readily be drawn (e.g., using novel problems, providing more detailed instructions, limiting the amount of response time available, and presenting participants with a concurrent working memory load should all have an impact on reasoning accuracy and the degree of belief bias). For instance, DPT predicts that limiting the amount of response time available should reduce logical reasoning accuracy and increase belief bias. Evans and Curtis-Holmes (2005) demonstrated exactly this: limiting the amount of time available to respond resulted in less accuracy and more belief bias as response bias (motivated reasoning was also eliminated, but it is not the main focus of the standard DPT of reasoning). This finding was taken as strong evidence in favour of the DPT of reasoning. As we have shown in Experiments 6 and 8, however, these results need to be nuanced: when the proper measurement techniques were used and individual differences were taken into account, it turned out that belief bias as response bias remained unaffected by a response time limit.

This raises the question whether the other manipulations which are often taken in support of DPT and though to be well understood stand up to scrutiny when the appropriate methodology is applied (e.g., SDT and individual differences). For instance, a concurrent working memory load has been shown to reduce reasoning accuracy and increase belief bias (De Neys, 2006). Likewise, strong deductive instructions have been shown to increase reasoning accuracy and eliminate or reduce belief bias (Newstead et al., 1992; Evans et al., 1994). Our finding that Evans and Curtis-Holmes' (2005) response time limit result was not replicated using the appropriate method suggests that similar replications should be conducted for the working memory load and instructions manipulations. This is especially important given the strong emphasis of DPT on belief bias as a means of justifying the appropriateness of the theory. If it turns out that the DPT account of belief bias does not hold up, then this may suggest that a uni process theory is sufficient to account for belief bias (Kruglanski, 2013). Such a finding would follow a trend in other domains of cognitive psychology such as memory (Berry, Shanks, & Henson, 2008) and learning (Mitchell, De

Houwer, & Lovibond, 2009), where recent research suggests that DPT might not be fundamental to explaining the current empirical findings.

5.3.4 Belief Bias in Other Arguments

One question which was not explored in this dissertation is whether the present results extend to other argument types. We focused exclusively on syllogistic reasoning, but belief bias is not unique to syllogisms. In the study of reasoning with causal conditionals, for instance, it has been demonstrated that response bias is pervasive unless participants are of higher cognitive ability and are given strong deductive reasoning instructions (Evans et al., 2010). Belief bias as response bias has also been demonstrated to occur in transitive inference, which is a different class of relational reasoning problems focusing on magnitude relations between categories (e.g., assuming that mice are taller than blubs and that blubs are taller than elephants, does it follow that mice are taller than elephants?; Andews, 2010). Belief bias is also present in informal reasoning tasks: participants are influenced by their prior beliefs when estimating the height of people (Sá, West, & Stanovich, 1999) and when evaluating the quality of various arguments (Thompson & Evans, 2012). Importantly, evidence of motivated reasoning has not been found for any of these problems, suggesting that motivated reasoning is a phenomenon unique to syllogistic reasoning. This seems odd, given the pervasiveness of motivated reasoning in daily life, for instance in climate change deniers (Lewandowsky, Cook, Oberauer, & Hubble-Marriott, 2013) and scientists (Francis, 2013). Perhaps it is not the case that motivated reasoning does not exist for these types of reasoning problems, but rather that the tasks have not been investigated using the appropriate methods. A reanalysis of such studies using a more appropriate measurement method such as SDT which also takes individual differences into account might lead to a more consistent picture about the nature of belief bias across experimental tasks.

5.3.5 Individual Differences in Memory

In contrast to reasoning research, the use of appropriate measurement methods has been a topic of much greater interest in research on memory. For instance, the shape of ROC curves has been taken as evidence in favour of (Yonelinas & Parks, 2007) or against (Wixted, 2007) the DPT of recognition memory. However, in spite of this strong focus on modelling, the debate on whether two types of processes (i.e., familiarity and recollection) are required to accommodate for recognition memory performance has not been resolved to this day (Jang, Wixted, & Huber, 2009; Pazzaglia, Dube, & Rotello, in press; Province & Rouder, 2012). We propose that taking individual differences into account while using the appropriate measurement method might be a fruitful strategy which could advance the debate in the field of recognition memory, just as it did for belief bias. For instance, our proposed individual differences model of belief bias could readily be extended to recognition memory in a source memory task (see Figure 5.2).

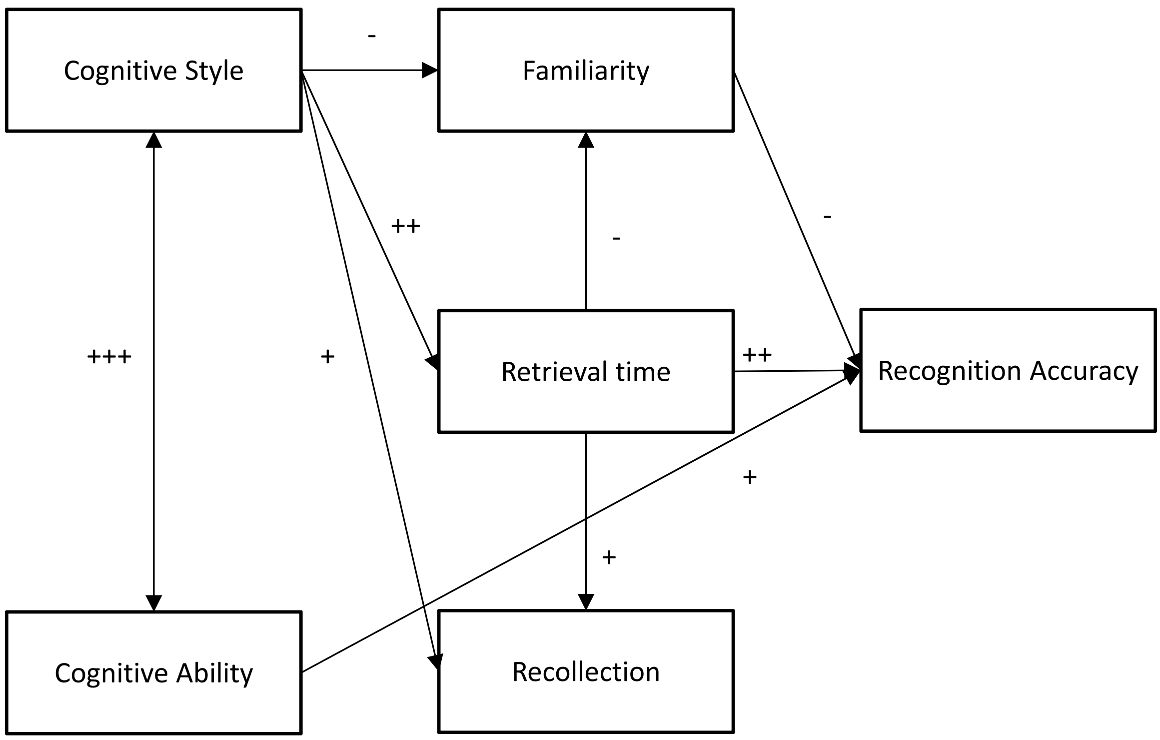


Figure 5.2. Possible extension of the individual differences model of belief bias to the case of recognition memory.

In parallel to the individual differences model of belief bias, this instantiation of the model for recognition memory proposes that cognitive ability positively predicts overall recognition memory performance, perhaps through its link with working memory capacity. Cognitive style, in contrast, predicts the strategic choice to engage in more retrieval time, leading to a greater reliance on recollection, and less reliance on familiarity, resulting in greater overall accuracy (in a source memory task). As with our individual differences model of belief bias, perhaps this model (or a better fitting variant of the model) could be used as a framework to relate several algorithmic level theories of recognition memory (e.g., low ACS = familiarity based processing in a continuous SDT-like fashion: Wixted, 2007; High ACS = recollection based processing in a discrete MPT-like way: Province & Rouder, 2012). Additional insights into the underlying type of processing could then stem from research in which people are tested simultaneously on reasoning and memory tasks, investigating the underlying pattern of correlations. For instance, if DPT is correct and familiarity and response bias are both based on type 1 processing, whereas recollection and motivated reasoning are based on type 2 processing, then we should find that people who engage in response bias more also rely more on familiarity, whereas people who show more motivated reasoning are more likely to engage in recollection, and that these relationships are mediated by cognitive style and ability.

5.4 Conclusion

In this dissertation we focused on two components of belief bias in human reasoning: response bias and motivated reasoning. We used an SDT approach to advance a theoretical debate between proponents of traditional belief bias accounts, according to which belief bias consists of both response bias and motivated reasoning, and proponents of the response bias account, according to which belief bias is just a response bias. We determined that belief bias is not just a response bias – at least not all the time. In doing so, we developed a novel forced choice reasoning method which has the unique ability to tap into response bias and motivated reasoning without the additional requirements of statistical modelling. In addition to this, the method also turned out to be capable of fully eliminating both belief bias components. The major finding was that individual differences play a crucial role in belief bias and its relation to reasoning. This led to the proposal of an individual differences account which provided a framework capable of uniting various extant theories of belief bias. We hope that future research using this framework will be useful in furthering our understanding of belief bias, human reasoning, and cognitive psychology in general.

References

Andrews, G. (2010). Belief-based and analytic processing in transitive inference depends on premise integration difficulty. *Memory & Cognition*, 38(7), 928–940

Aminoff, E. M., Clewett, D., Freeman, S., Frithsen, A., Tipper, C., Johnson, A., . . ., Miller, M. B. (2012). Individual differences in shifting decision criterion: A recognition memory study. *Memory & Cognition*. doi:10.3758/s13421-012-0204-6

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. doi: 10.1027/1618-3169.53.1.77

Barrouillet, P., Gauffroy, C., & Lecas, J. (2008). Mental models and the suppositional account of conditionals. *Psychological Review*, 115(3), 760–771. doi: 10.1037/0033-295X.115.3.760

Begg, I., & Denny, J. P. (1969). Empirical reconciliation of atmosphere and conversion interpretations of reasoning errors. *Journal of Experimental Psychology*, 81, 351–354.

Benjamin, A. S., Diaz, M. L., & Wee, S. (2009). Signal detection with criterion noise Application to recognition memory. *Psychological Review*, 116, 84–115.

Berry, C. J., Shanks, D. R., & Henson, R.N. (2008). A Unitary Signal-Detection Model of Implicit and Explicit Memory. *Trends in Cognitive Sciences*, 12 (10), 367–373.

Bonner, C., & Newell, B.R. (2010). In conflict with ourselves? An investigation of heuristic and analytic processes in decision making. *Memory & Cognition*, 38, 186–196.

Braine, M. D. S., & O'Brien, D. R. (1991). A theory of If: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, 98, 182–203.

Braine, M. D. S., & O'Brien, D. P. (1998). *Mental Logic*. Mahwah, NJ: Erlbaum.

Braine, M. D. S., & Rumain, B. (1983). *Logical Reasoning*. In J. Flavell and E. Markman (eds.), Carmichuel's Manual of Child Psychology, 4th edition: Cognitive Development. New York:Wiley.

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology*, 98(4), 550–558.

Burgess, A. E. (1995). Comparison of receiver operating characteristic and forced choice observer performance measurement methods. *Medical Physics*, 22, 643–656. doi:10.1118/1.597576

Byrne, R. M. J., & Johnson-Laird, P. N. (1990). *Remembering conclusions we have inferred: What biases reveal.* In J. P. Caverni, J. M. Fabre & M. Gonzalez (Eds.), Cognitive Biases: Their Contribution for Understanding Human Cognitive Processes (Vol. 68). Amsterdam: North-Holland.

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

Capon, A., Handley, S. J., & Dennis, I. (2003). Working memory and reasoning: An individual differences perspective. *Thinking & Reasoning*, 9, 203–244.

Chapman, L. J., & Chapman, J. P. (1959). Atmosphere effect re-examined. *Journal of Experimental Psychology*, 58, 220–226.

Chater, N., & Oaksford, M. (1999). The probabilistic heuristics model of syllogistic reasoning. *Cognitive Psychology*, 38, 191–258.

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

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

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, 17, 428–433.

Deutsch, R., Gawronski, B., & Strack, F. (2006). At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology*, 91, 385–405.

Deutsch, R., Kordts-Freudinger, R., Gawronski, B., & Strack, F. (2009). Fast and fragile: A new look at the automaticity of negation processing. *Experimental Psychology*, 56, 434–446.

Dickstein, L. S. (1978). The effect of figure on syllogistic reasoning. *Memory and Cognition*, 6, 76–83.

Dickstein, L. S. (1980). Inference errors in deductive reasoning. *Bulletin of the Psvchonomic Society*, 16, 414–416.

Dickstein, L. S. (1980). The meaning of conversion in syllogistic reasoning. *Bulletin of the Psvchonomic Society*, 16, 414–416.

Dube, C., Rotello, C. M., & Heit, E. (2010). Assessing the belief bias effect with ROCs: It's a response bias effect. *Psychological Review*, 117(3), 831–63. doi:10.1037/a0019634

Dube, C., Rotello, C. M., & Heit, E. (2011). The belief bias effect is aptly named: A reply to Klauer and Kellen (2011). *Psychological Review*, 118, 155–163.

Elqayam, S. & Evans, St.B.T. (2011). Subtracting 'ought' from 'is': Descriptivism versus normativism in the study of the human thinking. *Behavioral and Brain Sciences*, 34 (5), 233–248.

Evans, J. St. B. T., & Over, D. (2012). Reasoning to and from belief: Deduction and induction are still distinct. *Thinking & Reasoning*. doi:10.1080/13546783.2012.745450

Evans, J. St. B. T., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual process theory of reasoning. *Thinking & Reasoning*, 11(4), 382–389. doi:10.1080/13546780542000005

Evans, J. St. B. T., & Lynch, J. S. (1973).Matching bias in the selection task. *British Journal of Psychology*, 64, 391–397.

Evans, J. St. B. T., & Pollard, P. (1990). *Belief bias and problem complexity in deductive reasoning.* In J. P. Caverni., J. M. Fabre, & M. Gonzales (Eds.). Cognitive biases (pp. 131–154). Amsterdam: North-Holland.

Evans, J. St. B. T. & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives in Psychological Science*, 8(3), 223–241. doi:10.1177/1745691612460685

Evans, J. St. B. T. & Stanovich, K. E. (2013). Theory and metatheory in the study of dual processing. *Perspectives in Psychological Science*, 8(3), 263–271. doi:10.1177/1745691613483774

Evans, J. St. B. T. (1972). Interpretation and matching bias in a reasoning task. *Quarterly Journal of Experimental Psychology*, 24, 193–199.

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. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement*. Hove: Psychology Press.

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

Evans, J. S. B. T. (2011). Dual-process theories of reasoning: Contemporary issues and developmental applications. *Developmental Review*, 31(2-3), 86–102. doi:10.1016/j.dr.2011.07.007

Evans, J. S. B. T. (2012). Questions and challenges for the new psychology of reasoning. *Thinking & Reasoning*, 18, 5–31. doi:10.1080/13546783.2011.637674

Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, 11, 295–306.

Evans, J. St. B. T., Handley, S. J., & Harper, C. N. J. (2001). Necessity, possibility and belief: A study of syllogistic reasoning. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 54(A), 935–958.

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(6), 1495–1513. doi:10.1037//0278-7393.25.6.1495

Evans, J. St. B. T., Newstead, S. E., Allen, J. L., & Pollard, P. (1994). Debiasing by instruction: The case of belief bias. *European Journal of Cognitive Psychology*, 6, 263–285.

Evans, J. St. B. T., Over, D. E., & Handley, S. J. (2005). Suppositions, extensionality, and conditionals: A critique of the mental model theory of Johnson-Laird and Byrne (2002). *Psychological Review*, 112, 1040–1052.

Fechner, W. (1860). *Elemente der Psychophysik*. Breitkopf und Härtel:Leipzig

Francis, G. (2013). Replication, statistical consistency, and publication bias. *Journal of Mathematical Psychology*. doi:10.1016/j.jmp.2013.02.003

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

Gable, S. L., Reis, H. T., & Downey, G. (2003). He said, she said: a quasi-signal detection analysis of daily interactions between close relationship partners. *Psychological Science*, 14, 100–105.

George, C. (1995). The endorsement of the premises: Assumption-based or belief-based reasoning. *British Journal of Psychology*, 86, 93–111.

Gigerenzer, G., & Brighton, H. (2009). Homo Heuristicus: Why Biased Minds Make Better Inferences. *Topics in Cognitive Science*, 1(1), 107–143. doi:10.1111/j.1756-8765.2008.01006.x

Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning & Memory*, 2(1), 21–31. doi:10.1037//0278-7393.2.1.21

Gilinsky, A. S., & Judd, B. B. (1994). Working memory and bias in reasoning across the life span. *Psychology and Aging*, 9, 356–371.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford, England:Wiley.

Handley, S. J., Capon, A., Beveridge, M., Dennis, I., & Evans, J. St. B. T. (2004). Working memory and inhibitory control in the development of children's reasoning. *Thinking & Reasoning*, 10, 175–196. doi: 10.1080/13546780442000051

Heim, A. (1970). *AH4 group test of general intelligence manual*. London:Nelson.

Heit, E., & Hayes, B. K. (2011). Predicting reasoning from memory. *Journal of Experimental Psychology: General*, 140, 76–101. doi:10.1037/a0021488

Heit, E., & Rotello, C. M. (2005). *Are there two kinds of reasoning?* In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), Proceedings of the Twenty-Seventh Annual Meeting of the

Cognitive Science Society (pp. 923–928). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Heit, E., & Rotello, C. M. (2010). Relations between inductive reasoning and deductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(3), 805–812. doi:10.1037/a0018784

Henle, M., & Michael, M. (1956). The influence of attitudes on syllogistic reasoning. *Journal of Social Psychology*, 44, 115–127.

Henle, M. (1962). On the relation between logic and thinking. *Psychological Review*, 69(4), 366–378.

Huang, M. H., & Hauser, R. M. (1998). *Trends in Black–White test score differentials: II. The WORDSUM vocabulary test*. In U. Neisser (Ed.), The rising curve: Long-term gains in IQ and related measures (pp. 303–332). Washington DC: American Psychological Association.

Jang, Y., Wixted, J. T., & Huber, D. E. (2009). Testing signal-detection models of yes/no and two-alternative forced-choice recognition memory. *Journal of Experimental Psychology: General*,138(2), 291–306. doi:10.1037/a0015525

Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition*, 16, 1–62.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hove, UK: Psychology Press.

Johnson-Laird, P. N., & Steedman, M. (1978). The psychology of syllogisms. *Cognitive Psychology*, 10(1), 64–99. doi:10.1016/0010-0285(78)90019-1

Johnson-Laird, P. N. (1975). *Models of deduction.* In R. J. Falmagne (Ed.), Reasoning representation and process in children and adults (pp. 7–54). Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge: Cambridge University Press.

Kahneman, D., & Frederick, S. (2002) *Representativeness revisited: Attribute substitution in intuitive judgement*. In Heuristics and Biases: The Psychology of Intuitive Judgement (Gilovich, T., et al., eds), pp. 49–81, Cambridge University Press

Kahneman, D. (2011) *Thinking, Fast and Slow*. Farrar, Strauss: Giroux

Kantner, J., & Lindsay, D. S. (2012). Response bias in recognition memory as a cognitive trait. *Memory & Cognition*, 40(8), 1163–1177. doi:10.3758/s13421-012-0226-0

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, 42(3), 627–633.

Khemlani, S., & Johnson-Laird, P. N. (2012). Theories of the syllogism: A meta-analysis. *Psychological Bulletin*, 138(3), 427–457. doi:10.1037/a0026841

Klauer, K. C., & Kellen, D. (2011). Assessing the belief bias effect with ROCs: Reply to Dube, Rotello, and Heit (2010). *Psychological Review*, 118, 155–164. doi:10.1037/a0020698

Klauer, K. C., & Singmann, H. (in press). Does Logic Feel Good? Testing for Intuitive Detection of Logicality in Syllogistic Reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Klauer, K. C., & Kellen, D. (2011). Assessing the belief bias effect with ROCs: Reply to Dube, Rotello, and Heit (2010). *Psychological Review*, 118, 155–164. doi:10.1037/a0020698

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review*, 107, 852–884. doi:10.1037/0033- 295X.107.4.852

Kruglanski, A. W. (2013). Only One? The Default Interventionist Perspective as a Unimodel Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science*, 8, 242–247. doi:10.1177/1745691613483477

Lewandowsky, S., Cook, J., Oberauer, K., & Hubble-Marriott, M. (2013). Recursive fury: conspiracist ideation in the blogosphere in response to research on conspiracist ideation. *Frontiers in Psychology*. doi:10.3389/fpsyg.2013.00073

Lohman, D. F., & Lahkin, J. M. (2009). *Reasoning and Intelligence*. In R. J. Sternberg and S. B. Kaufman (Eds.), Handbook of Intelligence (2nd ed.). New York: Cambridge University Press.

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

Luce, R. D. (1963). A threshold theory for simple detection experiments. *Psychological Review*, 70, 61–79.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.

Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition*, 17(1), 11–17.

Mercier, H., & Sperber, D. (2011) Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34, 57–111

Milligan, G. W. (1987). The use of the arcsine transformation in the analysis of variance. *Educational and Psychological Measurement*, 47, 563–573. doi:10.1177/001316448704700303

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, 32(2), 183–198; discussion 198–246. doi:10.1017/S0140525X09000855

Morley, N. J., Evans, J. S. B. T., & Handley, S. J. (2004). Belief bias and figural bias in syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, 57(4), 666–692. doi:10.1080/02724980343000440

Morsanyi, K., & Handley, S. J. (2012). Logic feels so good-I like it! Evidence for intuitive detection of logicality in syllogistic reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 596–616. doi:10.1037/a0026099

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

Newstead, S. E., Pollard, P., Evans, J. S., & Allen, J. L. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition*, 45(3), 257–284.

Oakhill, J., Johnson-Laird, P. N., & Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition*, 31, 117–140.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford, UK: Oxford University Press.

Oaksford, M., & Chater, N. (2009). Précis of Bayesian rationality: The probabilistic approach to human reasoning. *Behavioral and Brain Sciences*, 32(1), 69-84; discussion 85–120. doi:10.1017/S0140525X09000284

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, 76, 972–987.

Pazzaglia, A. M., Dube, C., & Rotello, C. M. (in press). A critical comparison of discrete-state and continuous models of recognition memory: Implications for recognition and beyond. *Psychological Bulletin*.

Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2013). Belief bias during reasoning among religious believers and skeptics. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-013-0394-3

Pennycook, G., Cheyne, J. A., Seli, P., Koehler, D. J., & Fugelsang, J. A. (2012). Analytic cognitive style predicts religious and paranormal belief. *Cognition*, 123(3), 335–346. doi:10.1016/j.cognition.2012.03.003

Peterson, W. W., Birdsall, T. G., & Fox, W. C. (1954). The theory of signal detectability. *Transactions of the IRE Professional Group on Information Theory*, PGIT-4, 171–212.

Piaget, J. (1953). *Logic and psychology*. University of Manchester Press.

Polk, T. A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review*, 102(3), 533–566. doi:10.1037/0033-295X.102.3.533

Province, J. M., & Rouder, J. N. (2012). Evidence for discrete-state processing in recognition memory. *Proceedings of the National Academy of Sciences*, 109, 14357–14362.

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

Revlin, R., Leirer, V, Yopp, H., & Yopp, R. (1980). The belief-bias effect in formal reasoning: The influence of knowledge on logic. *Memory & Cognition*, 3, 584–585

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*. Cambridge, MA: MIT Press.

Rotello, C. M., & Heit, E. (2009). Modeling the effects of argument length and validity on inductive and deductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1317–1330. doi:10.1037/a0016648

Sá, W. C., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, 91, 497–510. doi:10.1037//0022-0663.91.3.497

Schroyens, W., Schaeken, W., & Handley, S. J. (2003). In search of counter-examples: Deductive rationality in human reasoning. *Quarterly Journal of Experimental Psychology*, 56(A), 1129–1145.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine*, 127, 966–972.

Sells, S. B. (1936). The atmosphere effect: An experimental study of reasoning. *Archives of Psychology No. 200.*

Shikishima, C., Yamagata, S., Hiraishi, K., Sugimoto, Y., Murayama, K., & Ando, J. (2011). A simple syllogism-solving test: Empirical findings and implications for g research. *Intelligence*, 39, 89–99.

Shynkaruk, J. M., & Thompson, V. A. (2006). Confidence and accuracy in deductive reasoning. *Memory & Cognition*, 34(3), 619–632.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Stanovich, K. E. (2010). *Rationality and the reflective mind*. New York: Oxford University Press.

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

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645–726.

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

Stenning K., & van Lambalgen M. S. (2008) *Human Reasoning and Cognitive Science*. Cambridge, MA: MIT Press.

Stevenson, R. J., & Over, D. E. (1995). Deduction from uncertain premises. *Quarterly Journal of Experimental Psychology*, 48(A), 613–643.

Stupple, E. J. N., Ball, L. J., Evans, J. St. B. T., & Kamal-Smith, E. (2011). When logic and belief collide: Individual differences in reasoning times support a selective processing model. *Journal of Cognitive Psychology*, 23, 931–941.

Thompson, V. A., & Evans, J. St. B. T. (2012). Belief bias in informal reasoning. *Thinking & Reasoning*, 18, 278–310.

Thompson, V. A. (1996). Reasoning from false premises: The role of soundness in making logical deductions. *Canadian Journal of Experimental Psychology*, 50, 315–319.

Thompson, V. A., Newstead, S. E., & Morley, N. J. (2011). *Methodological and theoretical issues in belief bias: Implications for dual-process theories.* In K. I. Manktelow, D. E. Over, & S. Elqayam (Eds.), The science of reason: A festschrift for Jonathan St. B. T. Evans (pp. 309–338). Hove, UK: Psychology Press.

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, 63, 107–140.

Thompson, V. A., Striemer, C. L., Reikoff, R., Gunter, R. W., & Campbell, J. I. D. (2003). Syllogistic reasoning time: Disconfirmation disconfirmed. *Psychonomic Bulletin & Review*, 10, 184–189.

Torrens, D., Thompson, V. A., & Cramer, K. M. (1999). Individual differences and the belief bias effect: Mental models, logical necessity, and abstract reasoning. *Thinking & Reasoning*, 5, 1–28.

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

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. London: Batsford.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 60, 273–281.

Wetherick, N. E., & Gilhooly, K. J. (1995). “Atmosphere”, matching, and logic in syllogistic reasoning. *Current Psychology*, 14, 169–178.

Wilkins, M. C. (1928). The effect of changed material on the ability to do formal syllogistic reasoning. *Archives of Psychology, 102.*

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, 114(1), 152–176.

Woodworth, R. S., & Sells, S. B. (1935). An atmosphere effect in syllogistic reasoning. *Journal of Experimental Psychology*, 18, 451–460.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: a review. *Psychological Bulletin*, 133(5), 800–832. doi:10.1037/0033-2909.133.5.800

Appendices

Appendix A: Instructions, Materials, Briefs, and Debriefs

Experiments 1 – 4

Instructions

Additional instructions were presented on the computer screen prior to the experiment. These can be found in the method sections of the relevant experiments.

In this experiment we are interested in people's reasoning. In a typical reasoning experiment, you would be asked to indicate whether the conclusion (the sentence under the line) is valid, i.e. whether it follows logically from the premises that precede it. A valid conclusion is one that must be true, assuming the premises are true. For example:

All diamonds are blutos	(1st premise)
All blutos are expensive things	(2nd premise)
All diamonds are expensive things	(conclusion)

This conclusion is logically valid because it necessarily follows from the premises.
In this experiment you will be choosing between two similar reasoning problems. You will be presented with 48 problem slides. Every slide will show two reasoning problems that have the same conclusion, but different premises. One of the problems will have a valid conclusion and one will have an invalid conclusion. Either the left problem has a valid conclusion or the right problem has a valid conclusion. It is your task to choose which argument is the valid one.

All diamonds are blutos All blutos are expensive things	All expensive things are blutos All blutos are diamonds
All diamonds are expensive things	All diamonds are expensive things

Notice that the conclusion is identical for the left problem and the right problem. However, the conclusion is only logically *valid* for the left problem, because it necessarily follows from the premises. The conclusion is logically *invalid* for the right problem, because it doesn't necessarily follow from the premises. **It is your task to indicate whether the conclusion is valid for the left problem or whether conclusion is valid for the right problem.** Use your mouse to choose either the left or the right problem. Choosing the left problem implies that you think that the left argument is valid and that the right argument is invalid. Choosing the right problem implies that you think that the right argument is valid and that the left argument is invalid.
Every time you choose you will also be asked to indicate your confidence in your response. Please use all three levels of confidence during the experiment.

If you don't understand the instructions, ask the experimenter to explain them to you now. Otherwise, follow the instructions on the screen.

Materials

These materials were used in all experiments, with the exception of Experiment 5 in which simple syllogisms were also used.

Syllogistic structures

Valid		Invalid	
EI3_O2:	No A are B	EI3_O1:	No A are B
	Some B are C		Some B are C
	Some C are not A		Some A are not C
EI4_O2:	No B are A	EI4_O1:	No B are A
	Some B are C		Some B are C
	Some C are not A		Some A are not C
EI1_O2:	No A are B	EI1_O1:	No A are B
	Some B are C		Some B are C
	Some C are not A		Some A are not C
IE1_O1:	Some A are B	IE1_O2:	Some A are B
	No B are C		No B are C
	Some A are not C		Some C are not A
OA3_O1:	Some A are not B	OE3_O1:	Some A are not B
	All C are B		No C are B
	Some A are not C		Some A are not C
AO3_O2:	All A are B	EO3_O2:	No A are B
	Some C are not B		Some C are not B
	Some C are not A		Some C are not A
OA4_O2:	Some B are not A	OE4_O2:	Some B are not A
	All B are C		No B are C
	Some C are not A		Some C are not A
AO4_O1:	All B are A	EO4_O1:	No B are A
	Some B are not C		Some B are not C
	Some A are not C		Some A are not C

Note. A = "All X are Y"; E = "No X are Y"; I = "Some X are Y"; O = "Some X are not Y". The first digit indicates the figure (1 = AB-BC; 2 = BA-CB; 3 = AB-CB; 4 = BA-BC). The second digit indicates the conclusion direction (1 = A-C; 2 = C-A).

Item contents

Category	Members			
amphibians	frogs	salamanders	toads	newts
birds	parrots	sparrows	ducks	robins
boats	kayaks	canoes	yachts	speedboats
cars	BMWs	Volvos	Vauxhalls	Fiats
criminals	robbers	murderers	embezzlers	terrorists
furniture	desks	sofas	cupboards	bookcases
dogs	Spaniels	Labradors	Terriers	Dalmatians
drinks	beers	sodas	wines	whiskeys
fish	trout	salmons	cods	haddocks
fruits	prunes	peaches	apples	bananas
insects	bees	beetles	ants	spiders
reptiles	lizards	iguanas	snakes	crocodiles
tools	hammers	saws	spanners	shovels
trees	oaks	willows	pines	maples
vegetables	carrots	cabbages	parsnips	radishes
weapons	cannons	swords	guns	spears

Note. In Experiment 1 only the first two columns of members were used for the believable and unbelievable problem types, whereas the third column was used as the linking term for the neutral problems.

Linking terms

redes	fosks	pives	pields	decottions	sothods	renes	bunges
wasses	geets	swants	cronxes	firters	nickhomes	revoules	pinds
foins	chindles	soats	sonds	pumes	papes	trops	envenches
lebs	brops	stoges	crots	punties	stamuses	vennars	cortemns
weens	quinces	loaxes	stoals	curges	gruts	cosuors	wightes
punds	jubs	parfs	fises	hoons	tutches	brimbers	punes
cofts	spashes	fimps	brams	heets	piffures	burtes	queels
flamps	dathses	darms	vosts	trinnels	goples	boodings	veemers

Note. Pseudowords were generated using Wuggy (Keuleers & Brysbaert, 2008).

Brief

On this sheet you will find all the information necessary for you to be able to give informed consent to take part in this experiment. Please read it carefully. You can ask the experimenter any questions you may have.

In this task you will be presented with 48 times 2 reasoning problems. For every trial you will be asked to choose which of both has a valid conclusion (see instructions for details).

Please remember that you have the right to stop your participation at any time. Also, your data will be kept confidential and the only connection between the two tasks is a participant code to make sure you remain anonymous. It follows that the data-analysis will also be completely anonymous. You have the right to withdraw your data after the experiment. If you care to do so, it will be removed from the analysis.

If you understand all these of these things and if you agree to them please read and sign the informed consent form on the back of this page.

UNIVERSITY OF PLYMOUTH

FACULTY OF SCIENCE

School of Psychology

CONSENT TO PARICIPATE IN RESEARCH PROJECT

Researcher:

Dries Trippas

Topic:

Confidence, choice and reasoning

The aim of this research is to study the relationship between reasoning and confidence

As a participant in this study you will be asked to judge the validity of 48 times 2 reasoning problems on a computer screen. All further instructions will be provided on an instruction sheet and on screen.

Upon finishing the experiment you will receive a written debrief with detailed information about the experiment and contact details for more information. You are also welcome to ask any further questions to the experimenter during and after the experiment.

The objectives of this research have been explained to me.

I understand that I am free to withdraw from the research at any stage, and ask for my data to be destroyed if I wish.

I understand that my anonymity is guaranteed, unless I expressly state otherwise.

I understand that the Principal Investigator of this work will have attempted, as far as possible, to avoid any risks, and that safety and health risks will have been separately assessed by appropriate authorities (e.g. under COSSH regulations)

Under these circumstances, I agree to participate in the research.

Name: ……………………………………………………………………..

Signature: …………………………………………….. Date: ………………………..

Debrief

Thank you for participating in this study.

In this study we wanted to investigate belief bias. Belief bias is the tendency for people to accept or refute arguments based on conclusion believability. There are two common ways of explaining belief bias. One explanation is that people simply accept believable conclusions more than unbelievable conclusions without actually reasoning about the problems (e.g., Dube, Rotello & Heit, 2010). Another series of theories claims that conclusion believability somehow influences the reasoning process. For instance: if the conclusion is believable, people will just accept it. However, if the conclusion is unbelievable people will look at the premises and start looking for counterexamples to see if the conclusion necessarily follows or not (Evans, 2003).

In this experiment we attempted to compare both accounts of belief bias by forcing people to choose between a valid and an invalid argument. The catch was that even though across the entire experiment conclusions differed in believability, the two arguments presented next to each other never differed in conclusion believability, only in validity. As such, you as a participant had no means to accept or reject arguments based on conclusion believability because one had to be valid and one had to be invalid, thus forcing you to reason about the underlying structure (or just guess if reasoning was too hard).

If you have any further questions, or if you want to withdraw you data, please feel free to contact us.

Researcher:	Dries Trippas:	dries.trippas@plymouth.ac.uk
Supervisors:	Professor Simon Handley:	shandley@plymouth.ac.uk
	Dr. Michael Verde:	michael.verde@plymouth.ac.uk

References:

Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences, 7,* 454-459

Dube, Rotello & Heit. (2010). Assessing the belief bias effect with ROCS: It's a response bias effect. *Psychological Review, 117,* 831-863.

Experiment 5

Instructions

In this experiment you will be asked to judge the validity of conclusions of 64 reasoning problems. A conclusion can either be valid or invalid. In order to judge whether the conclusion of a certain reasoning problem is valid or not, you need to apply the rules of logic. A conclusion is valid if it necessarily follows from the premises. A conclusion is invalid if it doesn't necessarily follow from the premises. You also must consider both premises to be true, even if they contain nonsense terms such as "blutos". Consider the following reasoning problem as an example:

All diamonds are blutos	(1st premise)
All blutos are expensive things	(2nd premise)
_______________________________	
All diamonds are expensive things	(conclusion)

This conclusion is logically valid because it necessarily follows from the premises.

Before the actual experiment begins you will receive 4 practice trials to make sure you understand the instructions. You can respond "valid" by pressing the "s" key or "invalid" by pressing the "k" key. After each valid/invalid decision you will also have to indicate how confident you are in your response by pressing "1" (=not at all confident), "2" (=moderately confident), or "3" (=very confident). Please use all three levels of confidence during the experiment. It is very important that you understand these instructions before beginning with the actual experiment. If you don't, please tell the experimenter and he will explain them to you again.

After the practice trials, the actual experiment will begin. When you have completed all 64 trials, you will be shown a slide that indicates the end of the experiment. When you see this, tell the experimenter you're done and you will be debriefed.

Materials

The materials for the complex problems were identical to those used in Experiments 1 – 4. The following additional syllogistic structures were used in the simple condition:

Valid		Invalid	
AE1_E1:	All A are B No B are C No A are C	AE1_I1:	All A are B No B are C Some A are B
AE3_E1:	All A are B No C are B No A are C	AE3_I1:	All A are B No C are B Some A are B
EA3_E1:	No A are B All C are B No A are C	EA3_I1:	No A are B All C are B Some A are B
EA2_E1:	No B are A All C are B No A are C	EA2_I1:	No B are A All C are B Some A are B
IA1_I1:	Some A are B All B are C Some A are C	IA1_E1:	Some A are B All B are C No A are B
AI4_I1:	All B are A Some B are C Some A are C	AI4_E1:	All B are A Some B are C No A are B
IA4_I1:	Some B are A All C are A Some A are C	IA4_E1:	Some B are A All B are C No A are B
AI2_I1:	All B are A Some C are B Some A are C	AI2_E1:	All B are A Some C are B No A are B
AE1_E2:	All A are B No B are C No C are A	AE1_I2:	All A are B No B are C Some C are A
AE3_E2:	All A are B No C are B No C are A	AE3_I2:	All A are B No C are B Some C are A
EA3_E2:	No A are B All C are B No C are A	EA3_I2:	No A are B All C are B Some C are A
EA2_E2:	No B are A All C are B No C are A	EA2_I2:	No B are A All C are B Some C are A
IA1_I2:	Some A are B All B are C Some C are A	IA1_E2:	Some A are B All B are C No C are A
AI4_I2:	All B are A Some B are C Some C are A	AI4_E2:	All B are A Some B are A No C are A
IA4_I2:	Some B are A All B are C Some C are A	IA4_E2:	Some B are A All C are A No C are A
AI2_I2:	All B are A Some C are B Some C are A	AI2_E2:	All B are A Some C are B No C are A

Note. A = “All X are Y”; E = “No X are Y”; I = “Some X are Y”; O = “Some X are not Y”. The first digit indicates the figure (1 = AB-BC; 2 = BA-CB; 3 = AB-CB; 4 = BA-BC). The second digit indicates the conclusion direction (1 = A-C; 2 = C-A).

Brief

On this sheet you will find all the information necessary for you to be able to give informed consent to take part in this experiment. Please read it carefully. You can ask the experimenter any questions you may have.

In this task you will be presented with 64 reasoning problems. For every reasoning problem you will be asked to judge the validity of its conclusion (see instructions for details).

Please remember that you have the right to stop your participation at any time. Also, your data will be kept confidential and the only connection between the two tasks is a participant code to make sure you remain anonymous. It follows that the data-analysis will also be completely anonymous. You have the right to withdraw your data after the experiment. If you care to do so, it will be removed from the analysis.

If you understand all these of these things and if you agree to them please read and sign the informed consent form on the back of this page.

UNIVERSITY OF PLYMOUTH

FACULTY OF SCIENCE

School of Psychology

CONSENT TO PARICIPATE IN RESEARCH PROJECT

Researcher:

Dries Trippas

Topic:

Reasoning about syllogisms

The aim of this research is to study the relationship between reasoning and confidence

As a participant in this study you will be asked to judge the validity of 64 reasoning problems on a computer screen. All further instructions will be provided on an instruction sheet and on screen.

Upon finishing the experiment you will receive a written debriefing with detailed information about the experiment and contact details for more information. You are also welcome to ask any further questions to the experimenter during and after the experiment.

The objectives of this research have been explained to me.

I understand that I am free to withdraw from the research at any stage, and ask for my data to be destroyed if I wish.

I understand that my anonymity is guaranteed, unless I expressly state otherwise.

I understand that the Principal Investigator of this work will have attempted, as far as possible, to avoid any risks, and that safety and health risks will have been separately assessed by appropriate authorities (e.g. under COSSH regulations)

Under these circumstances, I agree to participate in the research.

Name: ………………………………………………………………………..

Signature: ....…………….. Date: ....…………..

Debrief

Thank you for participating in this study.

In this study we wanted to investigate the relation between reasoning about syllogisms and confidence. You were randomly assigned to one of two conditions. In one condition participants were presented with single model syllogisms. In the other condition they were presented with multiple model syllogisms. A single model syllogism is easier to solve than a multiple model syllogism.

Previous research has shown that when people are asked to reason about syllogisms with conclusions that vary in believability, a belief by logic interaction occurs. This suggests that when people are confronted with an unbelievable conclusion, they will be motivated to reason more and consequently give the correct response more often. A recent paper has contested this belief bias effect in terms of a motivated reasoning account (Dube, Rotello & Heit, 2010). They found that people will be more inclined to respond valid if the conclusion is believable, but that their reasoning performance is equal for believable and unbelievable conclusions.

In this experiment we attempted to test a prediction based on this account of belief bias: If accuracy doesn't change but the only difference is a shift in response bias, than we should find a similar shift in response bias (and lack of shift in accuracy) for single model syllogisms compared to multiple model syllogisms.

If you have any further questions, or if you want to withdraw you data, please feel free to contact us.

Researcher:	Dries Trippas:	dries.trippas@plymouth.ac.uk
Supervisors:	Professor Simon Handley:	shandley@plymouth.ac.uk
	Dr. Michael Verde:	michael.verde@plymouth.ac.uk

References

Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences, 7,* 454-459

Dube, Rotello & Heit. (2010). Assessing the belief bias effect with ROCS: It's a response bias effect. *Psychological Review, 117,* 831-863.

Experiment 6

Instructions

Note. These are the instructions for the speeded condition. The instructions for the self-paced condition are identical to those of Experiment 5.

In this experiment you will be asked to judge the validity of conclusions of 64 reasoning problems under a time limit. A conclusion can either be valid or invalid. In order to judge whether the conclusion of a certain reasoning problem is valid or not, you need to apply the rules of logic. A conclusion is valid if it necessarily follows from the premises. A conclusion is invalid if it doesn't necessarily follow from the premises. You also must consider both premises to be true, even if they contain nonsense terms such as "blutos". Consider the following reasoning problem as an example:

All diamonds are blutos	(1st premise)
All blutos are expensive things	(2nd premise)
________________________________	
All diamonds are expensive things	(conclusion)

This conclusion is logically valid because it necessarily follows from the premises.

Before the actual experiment begins you will receive 4 practice trials to make sure you understand the instructions. You can respond "valid" by pressing the "s" key or "invalid" by pressing the "k" key. A red timer at the top of the screen will show how much time you have left to make a response. After each valid/invalid decision you will also have to indicate how confident you are in your response by pressing "1" (=not at all confident), "2" (=moderately confident), or "3" (=very confident). Please use all three levels of confidence during the experiment. It is very important that you understand these instructions before beginning with the actual experiment. If you don't, please tell the experimenter and he will explain them to you again.

After the practice trials, the actual experiment will begin. When you have completed all 64 trials, you will be shown a slide that indicates the end of the experiment. When you see this, tell the experimenter you're done and you will be debriefed.

Materials and Brief

The materials and brief were identical to those used in Experiment 5.

Debrief

The debrief for the self-paced condition was identical to the one used in Experiment 5.

Thank you for participating in this study.

In this study we wanted to investigate the relation between reasoning about syllogisms and confidence under a time limit. You were randomly assigned to one of two conditions. In one condition participants had unlimited time. In the other condition participants had to reason under a 10 second time limit.

Previous research has shown that when people are asked to reason about syllogisms with conclusions that vary in believability, a belief by logic interaction occurs. This is known as belief bias (Evans, 2003). This suggests that when people are confronted with an unbelievable conclusion, they will be motivated to reason more and consequently give the correct response more often. A recent paper has contested this belief bias effect in terms of a motivated reasoning account (Dube, Rotello & Heit, 2010). They found that people will be more inclined to respond valid if the conclusion is believable, but that their reasoning performance is equal for believable and unbelievable conclusions.

In this experiment we attempted to test a prediction based on this account of belief bias: If the only difference between believable and unbelievable items is a shift in response bias, then we should find a larger shift in response for the time limit condition compared to the no time limit condition.

If you have any further questions, or if you want to withdraw you data, please feel free to contact us.

Researcher:	Dries Trippas:	dries.trippas@plymouth.ac.uk
Supervisors:	Professor Simon Handley:	shandley@plymouth.ac.uk
	Dr. Michael Verde:	michael.verde@plymouth.ac.uk

References

Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences, 7,* 454-459

Dube, Rotello & Heit. (2010). Assessing the belief bias effect with ROCS: It's a response bias effect. *Psychological Review, 117,* 831-863.

Experiment 7

Instructions

The instructions in the normal instructions condition were identical to those used in the complex condition of Experiment 5. The instructions below are for the pragmatic instructions condition.

In this experiment you will be asked to judge whether a piece of information below a line follows from some other information above a line. In doing so, it is important that you respond as quickly as possible while following your gut instinct. Simply go with the first answer that comes to mind. In other words, you should base the answer on your intuition.

Before the actual experiment begins you will receive 4 practice trials to make sure you understand the instructions. You can respond “follows” by pressing the “s” key or “doesn’t follow” by pressing the “k” key. After each valid/invalid decision you will also have to indicate how confident you are in your response by pressing “1” (=not at all confident), “2” (=moderately confident), or “3” (=very confident). Please use all three levels of confidence during the experiment. It is very important that you understand these instructions before beginning with the actual experiment. If you don’t, please tell the experimenter and he will explain them to you again.

After the practice trials, the actual experiment will begin. When you have completed all 64 trials, you will be shown a slide that indicates the end of the experiment. When you see this, tell the experimenter you’re done and you will be debriefed.

Materials and Brief

The materials and brief were identical to those used in the complex condition of Experiment 5

Debrief

Thank you for participating in this experiment. You completed 64 reasoning problems (more specifically, syllogisms). People tend to be influenced by their prior beliefs when judging whether a conclusion is valid or not (e.g., Evans 2003). The aim of this study was to investigate the effect of different instructions on this so-called belief bias. You were randomly assigned to either a group with weak instructions or to a group with normal instructions. We predicted that people in the weak instructions group would be more influenced by beliefs than people in the normal instructions group (Evans, Newstead, Allen, & Pollard, 1994).

If you have any further questions, or if you want to withdraw you data, please feel free to contact us.

Researcher:	Dries Trippas:	dries.trippas@plymouth.ac.uk
Supervisors:	Prof. Simon Handley:	shandley@plymouth.ac.uk
	Dr. Michael Verde:	mfverde@plymouth.ac.uk

References:

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., Newstead, S. E., Allen, J. L., & Pollard, P. (1994). Debiasing by instruction: The case of belief bias. *European Journal of Cognitive Psychology*, *6 (3)*, 263-285.

Experiment 8

Instructions

The instructions for the reasoning task (both speeded and self-paced) were identical to those used in Experiment 6. Prior to the AH4 test, the following instructions were presented.

The AH4 is a test of general intelligence that consists of two parts. The first part is a measure of crystallized intelligence (e.g., vocabulary and following instructions). The second part is a measure of fluid intelligence (e.g., spatial reasoning and completing sequences).

Both parts are preceded by ten practice trials. These do not count towards your total score, but they should be completed to familiarize yourself with the task. For the actual test you will get 10 minutes to complete Part I and 10 minutes to complete Part II. On your desk you should find a question book and a response sheet. Please do not write in the actual question book, but fill out the responses on the response sheet under the corresponding number.

Before you begin, please fill out the date, your name, your age, your gender, and what stage you are currently in. After this, please complete the practice trials for Part I. Once everybody has filled out the form and completed the practice trials we can begin with the timed part of the experiment. If you have any questions whatsoever, ask them during the practice trials.

Do not start solving the actual questions until explicitly told to do so!

Materials

The materials were identical to the ones used in the complex condition of Experiment 5.

Brief

On this sheet you will find all the information necessary for you to be able to give informed consent to take part in this experiment. Please read it carefully. You can ask the experimenter any questions you may have.

The first task is a measure of general intelligence which involves solving a number of problems known as the AH4. There are two parts to the AH and completing both will take about 20 minutes in total.

The second task is a reasoning task in which you will be presented with 64 reasoning problems. For every reasoning problem you will be asked to judge the validity of its conclusion. Both tasks will be explained in more detail before the start of the experiment. The experiment in its entirety will take no more than 1 hour.

Please remember that you have the right to stop your participation at any time. Also, your data will be kept confidential and the only connection between the two tasks is a participant code to make sure you remain anonymous. It follows that the data-analysis will also be completely anonymous. You have the right to withdraw your data after the experiment. If you care to do so, it will be removed from the analysis.

If you understand all these of these things and if you agree to them please read and sign the informed consent form on the back of this page.

PLYMOUTH UNVERSITY
FACULTY OF SCIENCE
School of Psychology

CONSENT TO PARICIPATE IN RESEARCH PROJECT

Researcher:

Dries Trippas

Topic:

Reasoning and Intelligence

The aim of this research is to study the relationship between reasoning and intelligence

As a participant in this study you will first be asked to take the AH4, a test of general intelligence. In a second task you will have to judge the validity of 64 reasoning problems on a computer screen. All further instructions will be provided before you start the actual experiment.

Upon finishing the experiment you will receive a written debriefing with detailed information about the experiment and contact details for more information. You are also welcome to ask any further questions to the experimenter during and after the experiment.

The objectives of this research have been explained to me. I understand that I am free to withdraw from the research at any stage, and ask for my data to be destroyed if I wish.

I understand that my anonymity is guaranteed, unless I expressly state otherwise.

I understand that the Principal Investigator of this work will have attempted, as far as possible, to avoid any risks, and that safety and health risks will have been separately assessed by appropriate authorities (e.g. under COSSH regulations)

Under these circumstances, I agree to participate in the research.

Name: ……………………………………………………………………..

Signature: ……………………………………….. Date: ………………………..

Debrief

Thank you for participating in this study.

In this study we wanted to investigate the relation between reasoning about syllogisms under a time limit and intelligence. You were randomly assigned to one of two conditions. In one condition participants had unlimited time. In the other condition participants had to reason under a 10 second time limit.

Previous research has shown that when people are asked to reason about syllogisms with conclusions of varying believability, a belief by logic interaction occurs. This is known as belief bias (Evans, 2003). This suggests that when people are confronted with an unbelievable conclusion, they will be motivated to reason more and consequently give the correct response more often. A recent paper has contested this belief bias effect in terms of a motivated reasoning account (Dube, Rotello & Heit, 2010). They found that people will be more inclined to respond valid if the conclusion is believable, but that their reasoning performance is equal for believable and unbelievable conclusions.

Previous research has shown that general intelligence as measured by the AH4 is related to the ability to follow instructions when reasoning about belief bias problems (Evans, Handley, Neilens, Bacon, & Over, 2010). In this experiment, we measured general intelligence because we suspected that showing belief bias might be positively related to general intelligence. We also wanted to investigate the effect of imposing a time limit on this potential relation between belief bias and intelligence.

If you have any further questions, or if you want to withdraw you data, please feel free to contact us.

Researcher:	Dries Trippas:	dries.trippas@plymouth.ac.uk
Supervisors:	Professor Simon Handley:	shandley@plymouth.ac.uk
	Dr. Michael Verde:	michael.verde@plymouth.ac.uk

References

Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences, 7,* 454-459

Dube, C., Rotello, C. M., & Heit., E. (2010). Assessing the belief bias effect with ROCS: It's a response bias effect. *Psychological Review, 117,* 831-863.

Evans, J. St. B. T., Handley, S., Neilens, H., Bacon, A. M., & Over, D. E. (2010). The influence of cognitive ability and instructional set on causal conditional inference. *Quarterly Journal of Experimental Psychology, 63,* 892–909.

Experiment 9

Instructions and Materials

The instructions were identical to the forced choice instructions used in Experiments 1 – 4.

The materials were identical to the ones used in Experiment 4.

Brief

On this sheet you will find all the information necessary for you to be able to give informed consent to take part in this experiment. Please read it carefully. You can ask the experimenter any questions you may have.

The first task is a measure of general intelligence which involves solving a number of problems known as the AH4. There are two parts to the AH and completing both will take about 20 minutes in total.

The second task is a reasoning task in which you will be presented with 64 reasoning problems. For every reasoning problem you will be asked to decide which of two reasoning problems is valid. Both tasks will be explained in more detail before the start of the experiment. The experiment in its entirety will take no more than 1 hour.

Please remember that you have the right to stop your participation at any time. Also, your data will be kept confidential and the only connection between the two tasks is a participant code to make sure you remain anonymous. It follows that the data-analysis will also be completely anonymous. You have the right to withdraw your data after the experiment. If you care to do so, it will be removed from the analysis.

If you understand all these of these things and if you agree to them please read and sign the informed consent form on the back of this page.

Debrief

Thank you for participating in this study.

In this study we wanted to investigate belief bias. Belief bias is the tendency for people to accept or refute arguments based on conclusion believability. There are two common ways of explaining belief bias. One explanation is that people simply accept believable conclusions more than unbelievable conclusions without actually reasoning differently about believable and unbelievable problems (Dube, Rotello & Heit, 2010). A more traditional explanation is that conclusion believability impacts and alters the actual reasoning process. For instance: if a conclusion is unbelievable reason better because they are more motivated to refute the argument (Evans, 2003).

In this experiment we attempted to compare both accounts of belief bias by forcing you to choose between a valid and an invalid argument. Sometimes both conclusions were believable, sometimes they were both unbelievable, and sometimes one was believable and one was unbelievable. Prior research using this method has not produced findings in line with the traditional or the new account. We hypothesised that this lack of belief bias was because the side by side presentation of the arguments focussed the attention on the structure of the arguments, lowering the usefulness of believability as a cue. In this study we attempted to increase the usefulness of beliefs as a cue to drive reasoning by only ever showing one problem at a time, reducing the possibility for participants to make a decision on the basis of superficial structural problem characteristics. Finally, we also attempted to investigate whether this process is mediated in any way by cognitive ability.

If you have any further questions, or if you want to withdraw you data, please feel free to contact us.

Researcher: Dries Trippas: dries.trippas@plymouth.ac.uk

Supervisors: Professor Simon Handley: shandley@plymouth.ac.uk

Dr Michael Verde: michael.verde@plymouth.ac.uk

References:
Dube, C., Rotello, C. M., & Heit, E. (2010). Assessing the belief bias effect with ROCs: It's a response bias effect. *Psychological Review, 117*, 831–863. doi:10.1037/a0019634

Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences, 7,* 454-459

Experiment 10

Instructions and Materials

The instructions were identical to the forced choice instructions used in Experiments 1 – 4.

The materials were identical to the ones used in Experiment 4.

Brief

On this sheet you will find all the information necessary for you to be able to give informed consent to take part in this experiment. Please read it carefully. You can ask the experimenter any questions you may have.

The first task is a forced choice reasoning task in which you will be presented with two reasoning problems side by side 64 times. For every trial you will be asked to choose which of both has a valid conclusion (see instructions for details).

The second task consists of three simple reasoning problems which you will be asked to solve.

Both tasks will be explained in more detail before the start of the experiment. The experiment in its entirety will take no more than 45 minutes.

Please remember that you have the right to stop your participation at any time. Also, your data will be kept confidential and the only connection between the two tasks is a participant code to make sure you remain anonymous. It follows that the data-analysis will also be completely anonymous. You have the right to withdraw your data after the experiment. If you care to do so, it will be removed from the analysis.

If you understand all these of these things and if you agree to them please read and sign the informed consent form on the back of this page.

UNIVERSITY OF PLYMOUTH

FACULTY OF SCIENCE

School of Psychology

CONSENT TO PARICIPATE IN RESEARCH PROJECT

Researcher:

Dries Trippas

Topic:

Forced Choice and Simple Reasoning

The aim of this research is to study the effect of cognitive ability on forced choice reasoning

As a participant in this study you will first be asked solve a forced choice reasoning task. Next, you will be asked to solve three simple reasoning problems. All further instructions will be provided on an instruction sheet and on the screen.

Upon finishing the experiment you will receive a written debrief with detailed information about the experiment and contact details for more information. You are also welcome to ask any further questions to the experimenter during and after the experiment.

The objectives of this research have been explained to me.

I understand that I am free to withdraw from the research at any stage, and ask for my data to be destroyed if I wish.

I understand that my anonymity is guaranteed, unless I expressly state otherwise.

I understand that the Principal Investigator of this work will have attempted, as far as possible, to avoid any risks, and that safety and health risks will have been separately assessed by appropriate authorities (e.g. under COSSH regulations)

Under these circumstances, I agree to participate in the research.

Name: ………………………………………………………………………..

Signature: ....…………….. Date: ....…………..

Debrief

Thank you for participating in this study.

In this study we wanted to investigate belief bias. Belief bias is the tendency for people to accept or refute arguments based on conclusion believability. There are two common ways of explaining belief bias. One explanation is that people simply accept believable conclusions more than unbelievable conclusions without actually reasoning differently about believable and unbelievable problems (Dube, Rotello & Heit, 2010). A more traditional explanation is that conclusion believability impacts and alters the actual reasoning process. For instance: if a conclusion is unbelievable people's reasoning is better because they are more motivated to refute it (Evans, 2003).

In this experiment we attempted to compare both accounts of belief bias by forcing you to choose between a valid and an invalid argument. Sometimes both conclusions were believable, sometimes they were both unbelievable, and sometimes one was believable and one was unbelievable. Prior research using this method has not produced findings in line with the traditional or the new account. We hypothesised that this lack of belief bias was because the side by side presentation of the arguments focussed the attention on the structure of the arguments, lowering the usefulness of believability as a cue. In this study we attempted to increase the usefulness of beliefs as a cue to drive reasoning by only ever showing one problem at a time, reducing the possibility for participants to make a decision on the basis of superficial structural problem characteristics. Finally, we also attempted to investigate whether this process is mediated in any way by cognitive reflection, which is the tendency to resist a first intuitive response in favour of a more difficult analytic one. This was measured using the three simple reasoning problems presented after the reasoning study.

If you have any further questions, or if you want to withdraw you data, please feel free to contact us.

Researcher:	Dries Trippas:	dries.trippas@plymouth.ac.uk
Supervisors:	Professor Simon Handley:	shandley@plymouth.ac.uk
	Dr Michael Verde:	michael.verde@plymouth.ac.uk

References:

Dube, C., Rotello, C. M., & Heit, E. (2010). Assessing the belief bias effect with ROCs: It's a response bias effect. *Psychological Review, 117*, 831–863. doi:10.1037/a0019634

Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences, 7,* 454-459

Experiment 11

Instructions

The instructions were identical to those used in the self-paced condition of Experiment 5. Participants were instructed for the base rate, ratio bias, numeracy, wordsum, and CRT tasks using on screen instructions taken from Pennycook et al. (2012).

Materials

The reasoning materials were identical to those used in all the previous experiments.

The following measures were used to measure analytic cognitive style (ACS) and cognitive ability (CA).

Analytic Cognitive Style measures

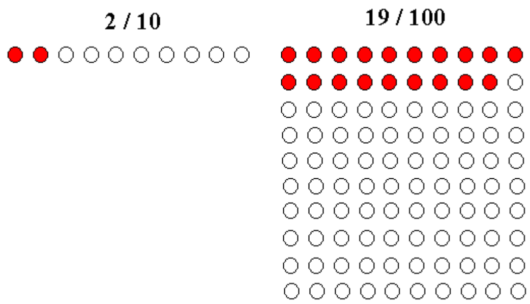
Base rate conflict (De Neys & Glumicic, 2008; Pennycook et al., 2012).

Example item:	In a study 1000 people were tested. Among the participants there were 5 engineers and 995 lawyers. Jack is a randomly chosen participant of this study. Jack is 36 years old. He is not married and is somewhat introverted. He likes to spend his free time reading science fiction and writing computer programs. What is most likely? a) Jack is a lawyer b) Jack is an engineer	"Analytic" answer = Jack is a lawyer "Intuitive" answer = Jack is an engineer

Cognitive reflection test (CRT; Frederick, 2005; Toplak, West, & Stanovich, 2011).

Example item:	A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?	Correct answer = 5 "Intuitive" answer = 10

Ratio bias (Bonner & Newell, 2010)

Example item:	**Suppose you were given two trays of jellybeans.** **If you were told that you would be paid $5 for choosing (without peeking) a red jellybean, from which tray would you select your jellybean?** 	"Analytic" answer = Left "Intuitive" answer = Right

Cognitive ability measures

Base rate neutral (De Neys & Glumicic, 2008; Pennycook et al., 2012)

Example item:	In a study 1000 people were tested. Among the participants there were 997 pool players and 3 basketball players. Jason is a randomly chosen participant of this study. Jason is 29 years old and has lived his whole life in New York. He has green coloured eyes and black hair. He drives a light-gray car. What is most likely? a) Jason is a pool player b) Jason is a basketball player	"Analytic" answer = Jason is a pool player No intuitive answer

Numeracy (Schwartz, Woloshin, Black, & Welch, 1997)

Example item:	In the BIG BUCKS LOTTERY, the chance of winning a $10 prize is 1%. What is your best guess about how many people would win a $10 prize if 1000 people each buy a single ticket to BIG BUCKS?	Correct answer = 10 No intuitive answer

WordSum (Huang & Hauser, 1998; Pennycook et al., 2012)

Example item:	Which word comes closest to the meaning of CAPRICE: a) value, b) a star, c) grimace, d) whim, e) inducement, f) don't know	Correct answer = whim No intuitive answer

Brief and Debrief

The brief and debrief were based on those used in Experiments 9 and 10, but adjusted for the fact that the experiment was run at the University of Waterloo, Canada, and that both cognitive ability and analytic cognitive style were tested.

Appendix B: ANOVA Tables

Experiment 1

The following analyses are presented in the order in which they are presented in the dissertation.

Accuracy

Motivated reasoning analysis

Effect	SS	df	MSE	F	p	η_p^2
Intercept	26.59	1	26.59	274.45	<.001	.94
Error(Intercept)	1.84	19				
Problem Type	0.04	1	0.04	0.58	.46	.03
Error(Problem Type)	1.24	19	0.07			

Note. Problem Type is a within subjects factor.

Full analysis

Effect	SS	df	MSE	F	p	η_p^2
Intercept	38.46	1	38.46	351.00	<.001	.95
Error(Intercept)	1.84	19				
Problem Type	0.06	2	0.03	0.71	.50	.04
Error(Problem Type)	1.71	38				

Note. Problem Type is a within subjects factor.

Confidence

Effect	SS	df	MSE	F	p	η_p^2
Intercept	255.23	1	255.23	839.96	<.001	.98
Error(Intercept)	5.77	19	0.30			
Problem Type	0.94	2	0.47	11.35	<.001	0.37
Error(Problem Type)	1.57	38	0.04			

Note. Problem Type is a within subjects factor.

Latency

Effect	SS	df	MSE	F	p	η_p^2
Intercept	5905.36	1	5905.36	38153.21	<.001	1
Error(Intercept)	2.94	19	0.16			
Problem Type	0.18	2	0.09	2.90	.068	.13
Error(Problem Type)	1.21	38	0.03			

Note. Problem Type is a within subjects factor.

Experiment 2

Accuracy

Response bias analysis

Effect	SS	df	MSE	F	p	η_p^2
Intercept	36.63	1	36.63	295.65	<.001	.91
Error(Intercept)	3.59	29	0.12			
Problem Type	0.001	1	0.001	0.02	.90	.001
Error(Problem Type)	2.41	29	0.08			

Note. Problem Type is a within subjects factor.

Motivated reasoning analysis

Effect	SS	df	MSE	F	p	η_p^2
Intercept	36.27	1	36.27	443.19	<.001	0.94
Error(Intercept)	2.37	29	0.08			
Problem Type	0.002	1	0.002	0.04	.84	.001
Error(Problem Type)	1.42	29	0.05			

Note. Problem Type is a within subjects factor.

Confidence

Effect	SS	df	MSE	F	p	η_p^2
Intercept	487.28	1	487.28	1339.96	<.001	.98
Error(Intercept)	10.55	29	0.36			
Problem Type	0.05	3	0.02	0.47	.70	.02
Error(Problem Type)	2.75	87	0.03			

Note. Problem Type is a within subjects factor.

Latency

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	11190.82	1	11190.82	26559.336	<.001	.99
Error(Intercept)	12.22	29	0.42			
Problem Type	0.08	3	0.03	1.37	.26	.05
Error(Problem Type)	1.61	87	0.02			

Note. Problem Type is a within subjects factor.

Experiment 3

Accuracy

Response bias analysis

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	27.49	1	27.49	296.19	<.001	.94
Error(Intercept)	1.86	20	0.09			
Problem Type	0.02	1	0.02	0.18	.68	.01
Error(Problem Type)	1.76	20	0.09			

Note. Problem Type is a within subjects factor.

Full analysis

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	42.61	1	42.61	336.56	<.001	.94
Error(Intercept)	2.53	20	0.13			
Problem Type	0.04	2	0.02	0.32	.73	.02
Error(Problem Type)	2.40	40	0.60			

Note. Problem Type is a within subjects factor.

Confidence

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	291.16	1	291.16	1107.76	<.001	.98
Error(Intercept)	5.26	20	0.26			
Problem Type	1.32	2	0.66	18.06	<.001	.48
Error(Problem Type)	1.46	40	0.04			

Note. Problem Type is a within subjects factor.

Latency

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	5892.99	1	2982.99	37224.07	<.001	.99
Error(Intercept)	3.22	20	0.16			
Problem Type	0.49	2	0.25	8.63	.001	.30
Error(Problem Type)	1.14	40	0.03			

Note. Problem Type is a within subjects factor.

Experiment 4

Accuracy

Response bias

Effect	SS	df	MSE	F	p	η_p^2
Intercept	86.42	1	86.42	742.30	<.001	.94
Error(Intercept)	5.94	51	0.12			
Problem Type	0.05	1	0.05	0.49	.49	.01
Error(Problem Type)	4.70	51	0.09			

Note. Problem Type is a within subjects factor.

Motivated reasoning

Effect	SS	df	MSE	F	p	η_p^2
Intercept	85.63	1	85.63	633.20	<.001	.93
Error(Intercept)	6.90	51				
Problem Type	0.27	1	0.27	4.10	.048	.074
Error(Problem Type)	3.37	51	0.07			

Note. Problem Type is a within subjects factor.

Confidence

Effect	SS	df	MSE	F	p	η_p^2
Intercept	1000.52	1	1000.52	1377.63	<.001	.964
Error(Intercept)	37.04	51	0.73			
Problem Type	0.37	3	0.12	3.33	.021	.061
Error(Problem Type)	5.62	153	0.04			

Note. Problem Type is a within subjects factor.

Latency

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	20335.14	1	20335.14	31930.38	<.001	.99
Error(Intercept)	32.48	51	0.64			
Problem Type	0.07	3	0.023	0.94	.43	.02
Error(Problem Type)	3.79	153	0.035			

Note. Problem Type is a within subjects factor.

Experiment 5

Endorsement

Effect	SS	df	MSE	F	p	η_p^2
Intercept	178.84	1	178.84	2114.22	<.001	.96
Complexity	1.07	1	1.07	12.69	.001	.13
Error(Intercept)	7.53	89	0.09			
Validity	51.09	1	51.09	398.66	<.001	.82
Validity x Complexity	12.03	1	12.03	93.86	<.001	.51
Error(Validity)	11.41	89	0.13			
Believability	1.07	1	1.07	10.08	.002	.10
Believability x Complexity	0.12	1	0.12	1.17	.28	.01
Error(Believability)	9.44	89	0.11			
Validity x Believability	0.19	1	0.19	5.47	.022	.058
Validity x Believability x Complexity	0.34	1	0.34	9.82	.002	.099
Error(Validity x Believability)	3.08	89	0.04			

Note. Complexity is a between-subjects factor. All other factors are manipulated within subjects.

SDT

Motivated reasoning

Effect	SS	df	MSE	F	p	η_p^2
Intercept	164.82	1	164.82	2201.00	<.001	.96
Complexity	7.24	1	7.24	96.68	<.001	.52
Error(Intercept)	6.67	89	0.075			
Believability	0.10	1	0.10	6.90	.010	.072
Believability x Complexity	0.083	1	0.083	5.51	.021	.058
Error(Believability)	1.35	89	0.015			

Note. Complexity is a between-subjects factor. Believability is a within subjects factor.

Response bias

Effect	SS	df	MSE	F	p	η_p^2
Intercept	4.16	1	4.16	26.76	<.001	.23
Complexity	10.18	1	10.18	65.44	<.001	.42
Error(Intercept)	13.84	89	0.16			
Believability	2.48	1	2.48	10.35	.002	.10
Believability x Complexity	.095	1	.095	0.40	.53	.004
Error(Believability)	21.34	89	0.24			

Note. Complexity is a between-subjects factor. Believability is a within subjects factor.

Latency

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	31080	1	31080	62837	<.001	.99
Complexity	1.30	1	1.30	2.64	.11	.03
Error(Intercept)	44.02	89	0.50			
Validity	0.03	1	0.03	1.94	.17	.02
Validity x Complexity	0.43	1	0.43	25.48	<.001	.22
Error(Validity)	1.50	89	0.02			
Believability	0.006	1	0.006	0.22	.64	.002
Believability x Complexity	0.002	1	0.002	0.052	.82	.001
Error(Believability)	2.59	89	0.029			
Validity x Believability	0.015	1	0.015	0.51	.48	.006
Validity x Believability x Complexity	0.027	1	0.027	0.91	.34	.010
Error(Validity x Believability)						

Note. Complexity is a between-subjects factor. All other factors are manipulated within subjects.

Experiment 6

Endorsement

Effect	SS	df	MSE	F	p	η_p^2
Intercept	152.18	1	152.18	1575.03	<.001	.95
Time Limit	0.23	1	0.23	2.35	.13	.027
Error(Intercept)	8.12	84	0.097			
Validity	5.82	1	5.82	67.26	<.001	.46
Validity x Time Limit	0.17	1	0.17	1.93	.17	.023
Error(Validity)	7.26	84	0.086			
Believability	3.66	1	3.66	30.01	<.001	.26
Believability x Time Limit	0.22	1	0.22	1.77	.19	0.21
Error(Believability)	10.25	84	.12			
Validity x Believability	0.27	1	0.27	4.39	.039	.050
Validity x Believability x Time Limit	0.048	1	0.048	0.78	.38	.009
Error(Validity x Believability)	5.19	84	0.062			

Note. Time Limit is a between-subjects factor. All other factors are manipulated within subjects.

SDT

Motivated reasoning

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	79.81	1	79.81	1295.69	<.001	.94
Time Limit	0.08	1	0.08	1.31	.26	.015
Error(Intercept)	5.17	84	0.062			
Believability	0.066	1	0.066	1.99	.16	.023
Believability x Time Limit	0.016	1	0.016	0.48	.49	.006
Error(Believability)	2.78	84	0.033			

Note. Time Limit is a between-subjects factor. Believability is a within subjects factor.

Response bias

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	8.48	1	8.48	44.68	<.001	.35
Time Limit	0.35	1	0.35	1.83	.18	0.21
Error(Intercept)	15.94	84	0.19			
Believability	7.65	1	7.65	32.81	<.001	.28
Believability x Time Limit	0.49	1	0.49	2.12	.15	.025
Error(Believability)	19.58	84	0.23			

Note. Time Limit is a between-subjects factor. Believability is a within subjects factor.

Latency

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	27935	1	27935	78693	<.001	.999
Time Limit	15.28	1	15.28	43.04	<.001	.34
Error(Intercept)	29.82	84	0.36			
Validity	0.49	1	0.49	28.02	<.001	.25
Validity x Time Limit	0.16	1	0.16	9.09	.003	.10
Error(Validity)	1.48	84	0.02			
Believability	0.009	1	0.009	0.36	.55	.004
Believability x Time Limit	0.002	1	0.002	0.09	.77	.001
Error(Believability)	2.20	84	0.026			
Validity x Believability	0.014	1	0.014	0.95	.33	.011
Validity x Believability x Time Limit	0.02	1	0.022	1.55	.22	.018
Error(Validity x Believability)	1.21	84	0.014			

Note. Time Limit is a between-subjects factor. All other factors are manipulated within subjects.

Experiment 7

Endorsement

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	129.31	1	129.31	873.39	<.001	.93
Instructions	0.091	1	0.091	0.62	.44	.01
Error(Intercept)	9.48	64	0.15			
Validity	6.53	1	6.53	47.22	<.001	.43
Validity x Instructions	0.25	1	0.25	1.81	.18	.028
Error(Validity)	8.85	64	0.14			
Believability	4.87	1	4.87	24.58	<.001	.28
Believability x Instructions	0.095	1	0.095	0.48	.49	.007
Error(Believability)	12.69	64	0.20			
Validity x Believability	0.002	1	0.002	0.029	.87	<.001
Validity x Believability x Instructions	0.005	1	0.005	0.093	.76	.001
Error(Validity x Believability)	3.57	64	0.056			

Note. Instruction is a between-subjects factor. All other factors are manipulated within subjects.

SDT

Motivated reasoning

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	69.28	1	69.28	926.46	<.001	.94
Instructions	0.10	1	0.10	1.39	.24	0.021
Error(Intercept)	4.79	64	0.075			
Believability	0.024	1	0.024	0.70	.41	0.011
Believability x Instructions	0.001	1	0.001	0.20	.89	<.001
Error(Believability)	2.20	64	0.034			

Note. Instruction is a between-subjects factor. Believability is a within subjects factor.

Response bias

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	8.96	1	8.96	30.19	<.001	.32
Instructions	0.20	1	0.20	0.66	.42	.01
Error(Intercept)	19.00	64	0.30			
Believability	10.16	1	10.16	25.53	<.001	.29
Believability x Instructions	0.12	1	0.12	0.311	.58	.005
Error(Believability)	25.47	64	0.40			

Note. Instruction is a between-subjects factor. Believability is a within subjects factor.

Latency

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	22415	1	22415	20977	<.001	.997
Instructions	2.47	1	2.47	2.31	.13	.035
Error(Intercept)	68.39	64	1.07			
Validity	0.54	1	0.54	25.022	<.001	.28
Validity x Instructions	0.035	1	0.035	1.65	.20	.025
Error(Validity)	1.38	64	0.022			
Believability	0.034	1	0.034	0.89	.35	.014
Believability x Instructions	.006	1	.006	0.17	.68	.003
Error(Believability)	2.43	64	0.038			
Validity x Believability	0.001	1	0.001	0.068	.80	.001
Validity x Believability x Instructions	0.005	1	0.005	0.219	.64	.003
Error(Validity x Believability)	1.35	64	0.021			

Note. Instruction is a between-subjects factor. All other factors are manipulated within subjects.

Experiment 8

Endorsement

Effect	SS	df	MSE	F	p	η_p^2
Intercept	144.83	1	144.83	880.15	<.001	.92
Time Limit	0.77	1	0.77	4.68	.034	.055
Cognitive Ability	0.09	1	0.09	0.55	.46	.007
Time Limit x Cognitive Ability	0.51	1	0.51	3.07	.084	.036
Error(Intercept)	13.33	81	0.17			
Validity	6.92	1	6.92	66.25	<.001	.45
Validity x Time Limit	1.61	1	1.61	15.37	<.001	.16
Validity x Cognitive Ability	0.36	1	0.36	3.46	.066	.041
Validity x Limit x Ability	0.001	1	0.001	0.013	0.91	<.001
Error(Validity)	8.46	81	0.104			
Believability	2.24	1	2.24	16.70	<.001	.17
Believability x Time Limit	0.059	1	0.059	0.44	.51	.005
Believability x Cognitive Ability	0.28	1	0.28	2.08	.15	.025
Believability x Limit x Ability	0.09	1	0.09	0.64	.43	.008
Error(Believability)	10.85	81	0.13			
Validity x Believability	0.05	1	0.05	0.87	.35	.011
Validity x Believability x Time Limit	.86	1	.86	14.92	<.001	.16
Validity x Believability x Ability	.016	1	.016	.28	.60	.003
Validity x Believability x Limit x Ability	0.024	1	0.024	0.42	.52	.005
Error(Validity x Believability)	4.70	81	.058			

Note. Time Limit and Cognitive Ability are between subjects factors. All other factors are manipulated within subjects.

SDT

Motivated reasoning

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	78.02	1	78.02	1133.90	<.001	.93
Time Limit	0.59	1	0.59	8.84	.005	.095
Cognitive Ability	0.33	1	0.33	4.81	.031	.056
Time Limit x Cognitive Ability	.002	1	.002	.03	.87	< .001
Error(Intercept)	5.57	81	0.07			
Believability	0.018	1	0.018	0.534	.47	.007
Believability x Time Limit	0.28	1	0.28	8.35	.005	0.93
Believability x Cognitive Ability	0.003	1	0.003	0.086	.77	.001
Believability x Limit x Ability	0.109	1	0.109	3.223	.076	.038
Error(Believability)	2.749	81	0.034			

Note. Time Limit and Cognitive Ability are between subjects factors.

Response bias

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	5.44	1	5.44	14.86	<.001	.16
Time Limit	0.94	1	0.94	2.56	.11	.031
Cognitive Ability	0.14	1	0.14	0.38	.54	.005
Time Limit x Cognitive Ability	1.05	1	1.05	2.87	.094	.034
Error(Intercept)	29.66	81	0.366			
Believability	5.26	1	5.26	19.00	<.001	.19
Believability x Time Limit	0.017	1	0.017	0.063	.80	.001
Believability x Cognitive Ability	0.69	1	0.69	2.50	.12	.03
Believability x Limit x Ability	0.13	1	0.13	0.47	.49	.006
Error(Believability)	22.42	81	0.28			

Note. Time Limit and Cognitive Ability are between subjects factors.

Latency

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	27322	1	27322	56631	<.001	.999
Time Limit	22.98	1	22.98	47.62	<.001	.37
Cognitive Ability	0.37	1	0.37	0.76	.39	.009
Time Limit x Cognitive Ability	0.003	1	0.003	0.006	.94	<.001
Error(Intercept)	39.08	81	.48			
Validity	0.62	1	0.62	16.31	<.001	.17
Validity x Time Limit	0.16	1	0.16	8.41	.005	.094
Validity x Cognitive Ability	0.006	1	0.006	0.32	.58	.004
Validity x Limit x Ability	0.001	1	0.001	0.034	.85	<.001
Error(Validity)	1.57	81	0.019			
Believability	0.008	1	0.008	0.342	.56	.004
Believability x Time Limit	0.003	1	0.003	0.140	.71	.002
Believability x Cognitive Ability	0.035	1	0.035	1.48	.23	.018
Believability x Limit x Ability	<.001	1	<.001	.003	.96	<.001
Error(Believability)	1.927	81	0.024			
Validity x Believability	.012	1	.012	0.65	.42	.008
Validity x Believability x Time Limit	0.002	1	0.002	0.13	.72	.002
Validity x Believability x Ability	< .001	1	<.001	0.001	.97	<.001
Validity x Believability x Limit x Ability	0.001	1	0.001	0.028	.87	<.001
Error(Validity x Believability)	1.442	81	0.018			

Note. Time Limit and Cognitive Ability are between subjects factors. All other factors are manipulated within subjects.

Experiment 9

Accuracy

Motivated reasoning

Effect	SS	df	MSE	F	p	η_p^2
Intercept	145.89	1	145.89	1225.09	<.001	.93
Cognitive Ability	1.79	1	1.79	15.02	<.001	.14
Error(Intercept)	11.19	94	0.12			
Problem Type	0.05	1	0.05	1.07	.30	.011
Problem Type x Cognitive Ability	.415	1	.415	9.58	.003	.092
Error(Believability)	4.075	94	0.043			

Note. Cognitive Ability is a between subjects factor. Problem Type is a within subjects factor.

Response bias

Effect	SS	df	MSE	F	p	η_p^2
Intercept	129.57	1	129.57	1488.59	<.001	.94
Cognitive Ability	1.52	1	1.52	17.21	<.001	.16
Error(Intercept)	8.29	94	0.09			
Problem Type	0.492	1	0.492	7.67	.007	.08
Problem Type x Cognitive Ability	0.235	1	0.235	3.66	.059	.04
Error(Believability)	6.03	94	.064			

Note. Cognitive Ability is a between subjects factor. Problem Type is a within subjects factor.

Confidence

Motivated reasoning

Effect	SS	df	MSE	F	p	η_p^2
Intercept	820.29	1	820.29	1904.73	<.001	.95
Cognitive Ability	1.477	1	1.47	3.41	.068	.035
Error(Intercept)	40.48	94	0.43			
Problem Type	0.024	1	0.024	0.73	.40	.008
Problem Type x Cognitive Ability	0.096	1	0.096	2.86	.094	.03
Error(Believability)	3.152	94	0.034			

Note. Cognitive Ability is a between subjects factor. Problem Type is a within subjects factor.

Response bias

Effect	SS	df	MSE	F	p	η_p^2
Intercept	851.64	1	851.64	2074.43	<.001	.96
Cognitive Ability	1.96	1	1.96	4.77	.031	.048
Error(Intercept)	38.59	94	0.41			
Problem Type	0.013	1	0.013	0.49	.49	.005
Problem Type x Cognitive Ability	0.048	1	0.048	1.83	.18	.019
Error(Believability)	2.44	94	0.026			

Note. Cognitive Ability is a between subjects factor. Problem Type is a within subjects factor.

Latency

Motivated reasoning

Effect	SS	df	MSE	F	p	η_p^2
Intercept	18589	1	18589	118185	<.001	.999
Cognitive Ability	0.001	1	0.001	0.008	.93	<.001
Error(Intercept)	14.79	94	0.16			
Problem Type	0.113	1	0.113	4.82	.031	.049
Problem Type x Cognitive Ability	0.001	1	0.001	0.05	.82	.001
Error(Believability)	2.20	94	0.023			

Note. Cognitive Ability is a between subjects factor. Problem Type is a within subjects factor.

Response bias

Effect	SS	df	MSE	F	p	η_p^2
Intercept	18326	1	18326	77287	<.001	.999
Cognitive Ability	0.12	1	0.12	0.50	.48	.005
Error(Intercept)	22.29	94	0.24			
Problem Type	0.009	1	0.009	0.363	.55	.004
Problem Type x Cognitive Ability	0.049	1	0.049	1.74	.19	.018
Error(Believability)	2.46	94	.026			

Note. Cognitive Ability is a between subjects factor. Problem Type is a within subjects factor.

Experiment 10

Accuracy

Motivated reasoning

Effect	SS	df	MSE	F	p	η_p^2
Intercept	85.50	1	58.50	636.19	<.001	.92
Block Order	0.089	1	0.089	0.66	.42	.01
Cognitive Style	0.37	1	0.37	2.75	.10	.045
Block Order x Cognitive Style	<.001	1	<.001	0.002	.96	<.001
Error(Intercept)	7.80	58	.13			
Problem Type	0.088	1	0.088	2.62	.11	.043
Problem Type x Block Order	0.064	1	0.064	1.90	.17	.032
Problem Type x Cognitive Style	0.421	1	0.421	12.48	.001	.177
Problem Type x Order x Style	0.003	1	0.003	0.094	.76	.002
Error(Problem Type)	1.96	58	0.034			

Note. Block Order and Cognitive Style are between subjects factors. Problem Type is a within subjects factor.

Response bias

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	81.77	1	81.77	918.56	<.001	.94
Block Order	0.514	1	0.514	5.77	.019	.091
Cognitive Style	0.605	1	0.605	6.80	.012	.105
Block Order x Cognitive Style	0.001	1	0.001	0.007	.93	<.001
Error(Intercept)	5.16	58	0.089			
Problem Type	1.249	1	1.249	11.17	.001	.16
Problem Type x Block Order	0.500	1	0.500	4.47	.039	.072
Problem Type x Cognitive Style	0.528	1	0.528	4.73	.034	.075
Problem Type x Order x Style	0.214	1	0.214	1.91	.17	.032
Error(Problem Type)	6.49	58	0.112			

Note. Block Order and Cognitive Style are between subjects factors. Problem Type is a within subjects factor.

Confidence

Motivated reasoning

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	453.49	1	453.49	1275.28	<.001	.96
Block Order	1.768	1	1.768	4.97	.030	.079
Cognitive Style	1.224	1	1.224	3.44	.069	.056
Block Order x Cognitive Style	0.003	1	0.003	0.009	.93	<.001
Error(Intercept)	20.63	58	0.356			
Problem Type	0.061	1	0.061	1.074	.30	.018
Problem Type x Block Order	0.031	1	0.031	0.541	.47	.009
Problem Type x Cognitive Style	0.189	1	0.189	3.31	.074	.054
Problem Type x Order x Style	0.132	1	0.132	2.31	.13	.038
Error(Problem Type)	3.31	58	0.057			

Note. Block Order and Cognitive Style are between subjects factors. Problem Type is a within subjects factor.

Response bias

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	488.18	1	448.18	1345.91	<.001	.96
Block Order	0.562	1	0.562	1.55	.22	.026
Cognitive Style	0.547	1	0.547	1.51	.22	.025
Block Order x Cognitive Style	0.022	1	0.022	0.061	.81	.001
Error(Intercept)	21.04	58	0.363			
Problem Type	0.125	1	0.125	4.25	.044	.068
Problem Type x Block Order	0.008	1	0.008	0.281	.60	.005
Problem Type x Cognitive Style	0.263	1	0.263	8.91	.004	.13
Problem Type x Order x Style	0.092	1	0.092	3.12	.083	.051
Error(Problem Type)	1.71	58	0.029			

Note. Block Order and Cognitive Style are between subjects factors. Problem Type is a within subjects factor.

Latency

Motivated reasoning

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	11020	1	11020	37360	<.001	.998
Block Order	4.265	1	4.265	14.46	<.001	.20
Cognitive Style	0.050	1	0.050	0.17	.68	.003
Block Order x Cognitive Style	0.024	1	0.024	0.082	.78	.001
Error(Intercept)	17.11	58	0.295			
Problem Type	0.007	1	0.007	0.332	.57	.006
Problem Type x Block Order	0.003	1	0.003	0.148	.70	.003
Problem Type x Cognitive Style	0.034	1	0.034	1.72	.20	.029
Problem Type x Order x Style	0.013	1	0.013	0.67	.42	.011
Error(Problem Type)	1.138	58	0.20			

Note. Block Order and Cognitive Style are between subjects factors. Problem Type is a within subjects factor.

Response bias

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	10866	1	10866	45724	<.001	.999
Block Order	1.968	1	1.968	8.281	.006	.125
Cognitive Style	0.812	1	0.812	3.418	.070	.056
Block Order x Cognitive Style	0.210	1	0.210	0.882	.352	.015
Error(Intercept)	13.784	58	0.238			
Problem Type	0.049	1	0.049	1.935	.169	.032
Problem Type x Block Order	0.006	1	0.006	0.229	.634	.004
Problem Type x Cognitive Style	0.029	1	0.029	1.158	.286	.020
Problem Type x Order x Style	0.004	1	0.004	0.150	.70	.003
Error(Problem Type)	1.473	58	0.025			

Note. Block Order and Cognitive Style are between subjects factors. Problem Type is a within subjects factor.

Experiment 11

Endorsement

Cognitive ability

Effect	SS	df	MSE	F	p	η_p^2
Intercept	421.47	1	421.47	2665.87	<.001	.93
Cognitive Ability	0.039	1	0.039	0.245	.621	.001
Error(Intercept)	29.723	188	0.158			
Validity	26.972	1	26.972	201.080	<.001	.517
Validity x Cognitive Ability	2.327	1	2.327	17.345	<.001	.084
Error(Validity)	25.217	188	0.134			
Believability	24.705	1	24.705	106.144	<.001	.361
Believability x Cognitive Ability	1.580	1	1.580	6.786	.010	.035
Error(Believability)	43.758	188	0.233			
Validity x Believability	0.293	1	0.293	5.538	.020	.029
Validity x Believability x Ability	0.495	1	0.495	9.367	.003	.047
Error(Validity x Believability)	9.943	188	0.053			

Note. Cognitive Ability is a between subjects factor. Validity and Believability are within subjects factors.

Analytic cognitive style

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	420.55	1	420.55	2657.65	<.001	.934
Cognitive Style	0.012	1	0.012	0.078	.78	<.001
Error(Intercept)	29.749	188	0.158			
Validity	26.008	1	26.008	196.78	<.001	.511
Validity x Cognitive Style	2.696	1	2.696	20.40	<.001	.098
Error(Validity)	24.848	188	0.132			
Believability	25.854	1	25.854	121.34	<.001	.392
Believability x Cognitive Style	5.295	1	5.295	24.86	<.001	.117
Error(Believability)	40.042	188	0.213			
Validity x Believability	0.243	1	0.243	4.45	.031	.025
Validity x Believability x Style	0.808	1	0.808	15.78	<.001	.077
Error(Validity x Believability)	9.630	188	0.051			

Note. Cognitive Style is a between subjects factor. Validity and Believability are within subjects factors.

SDT

Cognitive ability

Motivated reasoning

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	220.37	1	220.37	2268.74	<.001	.923
Cognitive Ability	2.572	1	2.572	26.476	<.001	.123
Error(Intercept)	18.261	188	0.097			
Believability	0.143	1	0.143	4.494	.035	.023
Believability x Cognitive Ability	0.284	1	0.284	8.909	.003	.045
Error(Believability)	5.985	188	0.032			

Note. Cognitive Ability is a between subjects factor. Believability is a within subjects factor.

Response bias

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	38.00	1	38.00	121.92	<.001	.393
Cognitive Ability	0.008	1	0.008	0.026	.872	<.001
Error(Intercept)	58.605	188	0.312			
Believability	50.493	1	50.493	108.227	<.001	.365
Believability x Cognitive Ability	2.449	1	2.449	5.249	.023	.027
Error(Believability)	87.710	188	0.467			

Cognitive Ability is a between subjects factor. Believability is a within subjects factor.

Analytic cognitive style

Motivated reasoning

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	217.177	1	217.177	2257.70	<.001	.923
Cognitive Style	2.748	1	2.748	28.566	<.001	.132
Error(Intercept)	18.084	188	0.096			
Believability	0.117	1	0.117	3.785	.053	.020
Believability x Cognitive Style	0.464	1	0.464	15.020	<.001	.074
Error(Believability)	5.805	188	0.031			

Note. Cognitive Style is a between subjects factor. Believability is a within subjects factor.

Response bias

Effect	**SS**	**df**	**MSE**	**F**	**p**	**η_p^2**
Intercept	37.885	1	37.885	121.515	<.001	.393
Cognitive Style	0.001	1	0.001	0.002	.97	
Error(Intercept)	58.613	188	0.312			
Believability	52.657	1	52.657	122.434	<.001	.394
Believability x Cognitive Style	9.304	1	9.304	21.632	<.001	.103
Error(Believability)	80.856	188	0.430			

Note. Cognitive Style is a between subjects factor. Believability is a within subjects factor.

Latency

Cognitive ability

Effect	**SS**	**df**	**MSE**	**F**	**p**	η_p^2
Intercept	67979	1	67979	57829	<.001	.997
Cognitive Ability	9.545	1	9.545	8.120	.005	.041
Error(Intercept)	220.998	188	1.176			
Validity	5.315	1	5.315	104.301	<.001	.357
Validity x Cognitive Ability	0.338	1	0.338	6.641	.011	.034
Error(Validity)	9.579	188	0.051			
Believability	0.042	1	0.042	1.037	.310	.005
Believability x Cognitive Ability	0.100	1	0.100	2.474	.117	.013
Error(Believability)	7.600	188	0.040			
Validity x Believability	0.027	1	0.027	0.631	.428	.003
Validity x Believability x Ability	0.017	1	0.017	0.400	.528	.002
Error(Validity x Believability)	7.927	188	0.042			

Note. Cognitive Ability is a between subjects factor. Validity and Believability are within subjects factors.

Analytic cognitive style

Effect	SS	df	MSE	F	p	η_p^2
Intercept	67672	1	67672	60152	<.001	.997
Cognitive Style	19.041	1	19.041	16.925	<.001	.083
Error(Intercept)	211.502	188				
Validity	5.131	1	5.131	102.369	<.001	.353
Validity x Cognitive Style	0.495	1	0.495	9.884	.002	.050
Error(Validity)	9.422	188				
Believability	0.036	1	0.036	0.891	.346	.005
Believability x Cognitive Style	0.075	1	0.075	1.837	.177	.010
Error(Believability)	7.625	188	0.041			
Validity x Believability	0.025	1	0.025	0.595	.441	.003
Validity x Believability x Style	0.007	1	0.007	0.162	.687	.001
Error(Validity x Believability)	7.937	188	0.042			

Note. Cognitive Style is a between subjects factor. Validity and Believability are within subjects factors.

Mediation

Motivated reasoning

Analytic Cognitive Style (ACS) -> Cognitive Ability (CA) -> Motivated Reasoning Index (MRI)

Model 1: MRI ~ Intercept + ACS

Effect	**Coefficient**	**Standard Error**	**t**	**p**
Intercept	-0.151	0.062	-2.432	.016
ACS	0.299	0.092	3.256	.002

Note. R^2 = .056; $F(1, 178)$ = 10.6

Model 2: MRI ~ Intercept + ACS + CA

Effect	**Coefficient**	**Standard Error**	**t**	**p**
Intercept	-0.164	0.091	-1.814	.07
ACS	0.286	0.112	2.563	.011
CA	0.030	0.147	0.206	.84

Note. R^2 = .056; $F(2, 177)$ = 5.294.

Model 3: CA ~ Intercept + ACS

Effect	**Coefficient**	**Standard Error**	**t**	**p**
Intercept	0.446	0.032	14.156	<.001
CA	0.429	0.047	9.137	<.001

Note. R^2 = .319, $F(1, 178)$ = 83.48

Mediation: Indirect effect = 0.013; *SE* = 0.063; *Z* = 0.206; *N* = 180

Cognitive Ability (CA) -> Analytic Cognitive Style (ACS) -> Motivated Reasoning Index (MRI)

Model 1: MRI ~ Intercept + CA

Effect	Coefficient	Standard Error	t	p
Intercept	-0.134	0.091	-1.469	.15
CA	0.244	0.123	1.974	.0499

Note. R^2 = .021; $F(1, 178)$ = 3.898.

Model 2: MRI ~ Intercept + CA + ACS

Effect	Coefficient	Standard Error	t	p
Intercept	-0.164	0.091	-1.814	.07
CA	0.030	0.147	0.206	.84
ACS	0.286	0.112	2.563	.011

Note. R^2 = .056; $F(2, 177)$ = 5.294.

Model 3: ACS ~ Intercept + CA

Effect	Coefficient	Standard Error	t	p
Intercept	0.106	0.060	1.758	.08
ACS	0.745	0.081	9.137	<.001

Note. R^2 = .319, $F(1, 178)$ = 83.48

Mediation: Indirect effect = 0.213; *SE* = 0.086; *Z* = 2.468; *N* = 180

Response bias

Analytic Cognitive Style (ACS) -> Cognitive Ability (CA) -> Response Bias Index (RBI)

Model 1: RBI ~ Intercept + ACS

Effect	**Coefficient**	**Standard Error**	**t**	**p**
Intercept	1.676	0.218	7.707	<.001
ACS	-1.557	0.323	-4.822	<.001

Note. R^2 = .116; $F(1, 178)$ = 23.25

Model 2: RBI ~ Intercept + ACS + CA

Effect	**Coefficient**	**Standard Error**	**t**	**p**
Intercept	2.053	0.316	6.507	<.001
ACS	-1.195	0.389	-3.070	.003
CA	-0.843	0.513	-1.643	.102

Note. R^2 = .129; $F(2, 177)$ = 13.09.

Model 3: CA ~ Intercept + ACS

Effect	**Coefficient**	**Standard Error**	**t**	**p**
Intercept	0.446	0.032	14.156	<.001
CA	0.429	0.047	9.137	<.001

Note. R^2 = .319, $F(1, 178)$ = 83.48

Mediation: Indirect effect = -0.361; SE = 0.224; Z = -1.62; N = 180

Cognitive Ability (CA) -> Analytic Cognitive Style (ACS) -> Response Bias Index (RBI)

Model 1: RBI ~ Intercept + CA

Effect	**Coefficient**	**Standard Error**	**t**	**p**
Intercept	1.676	0.218	7.707	<.001
CA	-1.557	0.323	-4.822	<.001

Note. R^2 = .116; *F*(1, 178) = 23.25

Model 2: RBI ~ Intercept + CA + ACS

Effect	**Coefficient**	**Standard Error**	**t**	**p**
Intercept	2.053	0.316	6.507	<.001
CA	-0.843	0.513	-1.643	.102
ACS	-1.195	0.389	-3.070	.003

Note. R^2 = .129; *F*(2, 177) = 13.09.

Model 3: ACS ~ Intercept + CA

Effect	**Coefficient**	**Standard Error**	**t**	**p**
Intercept	0.106	0.060	1.758	.08
ACS	0.745	0.081	9.137	<.001

Note. R^2 = .319, *F*(1, 178) = 83.48

Mediation: Indirect effect = -0.890; *SE* = 0.306; *Z* = -2.91; *N* = 180

Path Analysis

Full model (see Figure 4.6): χ^2 (1) = 0.629, p = .428

Regression weights:

Effect	Estimate	SE	Z	p
RT ~ ACS	0.875	0.235	3.723	<.001
RT ~ CA	0.334	0.307	1.088	0.276
MRI ~ ACS	0.243	0.117	2.082	0.037
MRI ~ CA	0.013	0.147	0.088	0.93
RBI ~ ACS	-0.767	0.396	-1.94	0.052
RBI ~ CA	-0.664	0.498	-1.333	0.183
MRI ~ RT	0.075	0.036	2.072	0.038
RBI ~ RT	-0.514	0.123	-4.185	<.001
Az ~ MRI	0.005	0.057	0.09	0.928
Az ~ RBI	-0.06	0.017	-3.554	<.001
Az ~ CA	0.307	0.111	2.772	0.006
Az ~ ACS	0.13	0.089	1.451	0.147
Az ~ RT	0.143	0.029	4.98	<.001

Note. RT = response time, ACS = cognitive style, CA = cognitive ability, MRI = motivated reasoning index, RBI = response bias index, Az = reasoning accuracy

Reduced model (see Figure 4.7): χ^2 *(6)* = 5.66, *p* = .45.

Regression weights:

Effect	**Estimate**	**SE**	**Z**	**p**
RT ~ ACS	1.015	0.197	5.145	<.001
RBI ~ ACS	-1.033	0.344	-3.005	0.003
RBI ~ RT	-0.527	0.123	-4.287	<.001
MRI ~ ACS	0.249	0.101	2.461	0.014
Az ~ RBI	-0.063	0.016	-3.841	<.001
Az ~ CA	0.385	0.095	4.063	<.001
MRI ~ RT	0.075	0.036	2.086	0.037
Az ~ RT	0.153	0.028	5.482	<.001

Note. RT = response time, ACS = cognitive style, CA = cognitive ability, MRI = motivated reasoning index, RBI = response bias index, Az = reasoning accuracy