Lie Detection: Cognitive Processes

Christopher Norman Howell Street

Division of Psychology and Language Sciences
University College London, UK

A dissertation submitted in part fulfilment of the requirement for the degree of
Doctor of Philosophy in Psychology

Prepared under the supervision of
Dr. Daniel C. Richardson

November 2013

I, Christopher Norman Howell Street, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature

Abstract

How do we make decisions when we are uncertain? In more real-world settings there is often a vast array of information available to guide the decision, from an understanding of the social situation, to prior beliefs and experience, to information available in the current environment. Yet much of the research into uncertain decision-making has typically studied the process by isolating it from this rich source of information that decision-makers usually have available to them. This thesis takes a different approach.

To explore how decisions are made under uncertainty in more real-world settings, this thesis considers how raters decide if someone is lying or telling the truth. Because people are skilled liars, there is little information available to make a definitive decision. How do raters negotiate the ambiguous environment to reach a decision?

Raters show a truth bias, which is to say they judge statements as truthful more often than they are so. Recent research has begun to consider dual process theories, suggesting there are two routes for processing information. They claim the truth bias results from an error-prone processing route, but that a more effortful and analytical processing route may overcome it.

I will generate a set of testable hypotheses that arise from the dual process position and show that the theory does not stand up to the test. The truth bias can be better explained as resulting from a single process that attempts to make the most

informed guess despite being uncertain. To make the informed guess, raters come to rely on context-relevant information when the behaviour of the speaker is not sufficiently diagnostic.

An adaptive decision maker position is advocated. I propose the truth bias is an emergent property of making the best guess. That is, in a different context where speakers may be expected to lie, a bias towards disbelieving should be seen. I argue context-dependency is key to understanding decision-making under uncertainty.

Table of Contents

List of Figures

List of Tables

Dedication

- Phyllis Street -

You made the most of your time and enjoyed the social life. It reminds me that there is more to life than a career path. Your constant encouragement gave new meaning to this monograph, a densely packed volume likely to be read by few, and enjoyed by fewer. You have contributed to this work more than you would ever know. I wish I could give you a copy.

Acknowledgements

This volume was the work of many wonderful people, from those who kept me sane during my graduate years to those whose discussions and insightful comments shaped the direction of my research and this thesis.

First, to my parents Julie L'homme and Wayne Street. Their faith and encouragement has brought me this far, and I love you both for it. This acknowledgement cannot begin to express my gratitude. I know you will still be there for me in the years to come. I hope I have lived up to your expectations.

I have been lucky enough to have some wonderful friendships, each of which have brought me joy and kept me sane. To Jayne Roberts and Ruth Challoner, you have selflessly given and never asked anything in return. You have been both friend and parent to me over many years.

I have far too many people to thank, and I know I will miss many important friends from this list, to whom I apologise. In no particular order, David and Sarah Roberts, Gregor Murray, Myshi MacKenzie, Emily Webber, Gwyneira Edwards and Rhys Dobbs, thank you for your continued friendship and putting up with my odd ways. Additional thanks to my many friends here at UCL who number far too many to list individually. Special thanks to Laura Speed for proof reading two chapters of this thesis; I imagine it was tedious work. I would also like to thank Matthias Gobel for his interesting discussions on the more social aspects of this thesis.

My collaborators, past and present, have taught me much about experimental design and creative thinking. Thank you to Jaume Masip, Miguel Vadillo, Benedetto De Martino, Catrine Jacobsen, Sarah Jordan, Sam Baron, Lee Tbaily, Costi Rezlescu and many others. Thanks also to Maarten Speekenbrink, Richard Morey and Dave Vinson for your help with the statistical questions I have had. I am especially grateful for the friendship of Nick Duran and Rick Dale: despite my awkward ways I hope we will continue to collaborate for years to come. As a late addition to this section, my examining committee Arndt Bröder and Nigel Harvey have given me much food for thought and made important contributions to this thesis.

To those who have shielded me from the barrage of administration, I am sure I have more to thank you for than I realise. Thanks to the general office staff Pia Horbacki, Hannah Spikesley and Joanna Strange and to Steve Williams for his ever-friendly support. Special thanks to John Draper and Steve Davis who, amongst other things, were the ringmasters of the UCL Psychology Christmas Parties. To the IT team in the basement, and especially to Dave Hetherington Mac-extraordinaire, who nursed my laptop back to health over the years from the torturous beatings I have often given it. I am ever grateful.

My undergraduate supervisor Ben W Tatler supported me during my formative years, both intellectually and emotionally. Your lectures sparked excitement in me. You encouraged me to work in your laboratory when I was able, and I enjoyed it thoroughly, with the exception of extracting eye movement data from your mobile eye trackers. That was thoroughly tedious. You were a wonderful mentor and showed me how psychological science should be carried out. I am sure we will continue to cross paths and hopefully begin some new collaborations.

Finally, to Daniel C Richardson. You were daring enough to take me on as a research assistant in your lab, and had faith in me to take up a PhD position with you on a research topic that neither of us had explored before. It has been exciting and I have learnt so much from you. You approach research with a creativity that I aspire to one day reach. Most importantly, you gave me intellectual freedom and allowed me to take my own route down the rabbit-hole that is cross-disciplinary research. For that I thank you.

Collaborations

Chapter 4: Jaume Masip, University of Salamanca.

Chapter 6: Costi Rezlescu, Harvard University.

To my readers, I hope this thesis will be of perfect thickness as to prop up your table leg.

Summary of Experiments

A brief summary the aims and outcomes of each experiment conducted is presented here for reference.

Chapter 4: Testing the Dual Process Theory: The Multi-Response Interview

Experiment 1: The Behavioural Account. The truth bias was shown to have an independent cognitive component. However, processing duration available to the rater could not predict the decline in the proportion of truth judgments (PTJ), failing to support two classes of heuristic-analytic models (HAMs).

Experiment 2: Channel Effects. A third class of HAMs was examined, which proposes the cues available in the environment determine whether a heuristic or analytical process is used. No support for this account was found, showing a decrease in the PTJ across multiple judgments regardless of channel availability. In addition, the amount of processing time available was not able to predict the decrease in bias.

Experiment 3: Thin-Slicing. The amount of processing time available was manipulated experimentally. However, there was no support for the suggestion that processing time determines the extent of the decreased bias, failing to support HAMs.

Experiment 4: The Consistency Heuristic. An alternative account was sought: raters may be making use of consistency information. This finding was supported. Interestingly, this cue proved to be diagnostic, suggesting raters are attending to useful information. However, that raters did not generally become more accurate over time suggests there may have been use of other information in the belief forming process.

Chapter 5: The Adaptive Decision Maker

Experiment 5: A Cartesian Mind – Online Comprehension. The Spinozan mind account proposes that, in order to understand, it is necessary to first automatically accept all information to be true. I suggest that this initial bias towards believing actually reflects the use of prior knowledge when uncertain but forced into judgment. This experiment replicates the Spinozan effect of an early truth bias when forced to judge, but there is no evidence of an early bias when able to withhold judgment and indicate uncertainty.

Experiment 6: Most People Tell the Truth – The Availability Heuristic. When forced into judgment, what information is being used to guide the judgment? It is suggested that situation-relevant context knowledge can be used. In the case of lie detection, base rate information from prior experience can be brought to the judgment: people usually tell the truth. This experiment shows that, when forced into judgment, early on in the decision process the base rate information is taken into account.

Experiment 7: A Cartesian Mind – Post Comprehension. Informed Cartesian raters are thought to make use of context information under uncertainty. This experiment aimed to test this account at the point of the final judgment. Surprisingly, and in direct contrast to Experiment 5, the truth bias was greatest when participants were *not* forced to judge. Exploration of the data suggested that the unforced condition did not show an increase in truth responding, but rather a decrease in lie responding in favour of indicating uncertainty. This experiment led to a distinction being made between internal uncertainty, resulting from an inability to form a decision in spite of the evidence, and external uncertainty, resulting from a lack of information available in the environment. In the absence of information (external uncertainty), it is proposed raters make use of prior knowledge and contextually relevant information. When information is available but uncertainty remains, context-specific knowledge can again be employed: truths are typically more difficult to spot than lies, and so an inability to decide may be taken as an indication of deception.

Experiment 8: The Cartesian Hopi Word Experiment. In an attempt to replicate the findings of Experiment 5 and to conduct a confirmatory study to replicate the findings of Experiment 7, the Gilbert, Krull and Malone (1990, Study 1) experimental setup was used. Whereas the Spinozan mind predicts an early bias towards believing, the adaptive decision-maker proposes uncertainty initially and a truth bias later resulting from a shift away from lie responding towards unsure responding. That is, a crossover interaction was predicted. Whilst the data partially support the claim, it was found that the experimental design is not sufficiently placed to be able to test the claims of either the Spinozan mind or the adaptive decision-maker. Forgetting effects could better explain the results.

Chapter 6: Social Orientation Theory

Experiment 9: Socialisation Practices. The reduction in lie responses at the point of final judgment may not reflect true uncertainty but rather a strategy that aims to avoid the socially aggressive act of labelling another without voluntarily making an incorrect judgment – it may act as an 'out'. The ostrich effect and the accusatory reluctance account both claim at their core that the social practices we engage with make us reluctant to label another a liar. By removing the implied social presence, a reduction in the truth bias was predicted under this account. It was found that whether the speaker was considered a social or non-social agent did not influence the degree of bias.

Experiment 10: Social Relatedness as a Heuristic. It was noted that the findings of Experiment 9 appear to contradict prior research showing that we are more inclined to believe those to who we feel closer. This experiment sought to determine whether social relatedness information, in isolation of other information that tends to accompany it (e.g. familiarity), can itself be used as a guide to aid the decision process. Although the truth bias cannot be accounted for as an accusatory reluctance (Experiment 9), social information can nonetheless be incorporated into the judgment.

Chapter 1: Overview

You are looking to buy a lottery ticket. In the lottery, five numbers are drawn at random, each with equal likelihood of being drawn. A friend gives you the option of buying one of two tickets: either a ticket with (a) 1, 2, 3, 4 and 5 selected, or with (b) 4, 19, 6, 32, and 12 selected. Each ticket is equally likely to win, and so it may be expected that you would equally choose one ticket as the other. In studies similar to this setup, that is not what happens. Instead, across a variety of tasks people show a systematic bias towards choosing one option over the other (Chaiken & Trope, 1999; Gilovich, Griffin & Kahneman, 2002; Plott & Smith, 2008).

This thesis explores the underlying processes involved in biased responding. To date, much of the decision-making research has confined itself to low-level laboratory conditions in an attempt to isolate the processes from social, contextual and even memory influences. But consider the lottery example. There is typically a bias towards choosing numbers that show less systematicity and ordering, a variation on the gambler's fallacy (Beach & Swensson, 1967; Jarvick, 1951; Witte, 1964). Although the lottery has no memory, you do. One account suggests the bias or fallacy may be a reflection of habits learned from life outside the laboratory that raters brought with them to the task (Ayton, Hunt & Wright, 1989; Estes, 1964; Lopes & Oden, 1987).

The overarching aim of my thesis is to understand how raters make decisions when there is little information available in more naturalistic environments. The real world is a rich source of information, and from it we may learn not only habits, but also the dynamics of and the rule that govern social interactions, the outcomes from past decisions, and expectations for future, similar decisions. This thesis will consider a decision-making task that makes these sources of information relevant and potentially useful for incorporating into the judgment. I will end by concluding that raters attempt to make the best guess possible given the lack of reliable information by relying on context-relevant information, and that these best guesses are the cause of biased responding in high-level socially oriented tasks. My research highlights the need to consider decision-making in context.

To explore uncertain decision making in situ, I will make use of a real-world decision that people make: deciding whether others are lying or telling the truth. Lie detection is an inherently difficult decision when people are such good liars. The cues available are minimal at best (see Levine, 2010 for the suggestion that most liars are undetectable). For that reason, lie detection offers a perfect test bed for exploring natural decision-making under uncertainty. Despite the uncertainty in the environment, raters do not responding randomly: they show a truth bias. The truth bias has been reported across various studies and meta-analyses (e.g., Bond & DePaulo, 2006; Fiedler & Walka, 1993; Levine, Park & McCornack, 1999; Masip, Garrido & Herrero, 2006, 2009; Stiff, Kim & Ramesh, 1992; Vrij, 2008; Zuckerman, DeFrank, Hall, Larrance & Rosenthal, 1979). It is defined operationally as making truth judgments more often than truthful statements are presented (c.f. Fiedler & Walka, 1993).

I will explore the cognitive processes underlying the truth bias in this thesis. Rarely has the lie detection field considered the cognitive processes underlying the judgment, but one account that has recently received some support is the dual process model, outlined below. Although findings have been consistent with the dual process model, lie detection researchers have yet to provide a true test of its explanatory power. I will present a series of challenges against the dual process model that it must meet if it is to account for the truth bias. I will show that the model fails to meet any of the challenges issued. Instead, I will show how the lie-truth judgment can be better thought of as being implemented by a single process rather than by two distinct processing streams. The major conclusion of this thesis is that we are adaptive decision makers: I suggest the truth bias is not an inherent component of the cognitive system but rather is a response bias that emerges from an integration of the cues in the environment with context-relevant knowledge to make informed guesses when unsure.

A dual-process model has the advantage of explaining not only how the decision is formed but why it is biased, i.e. because one of the processing routes makes fast but error-prone decisions. If it is shown that we are single-process decision makers, which my research suggests, then we lose an understanding of why the system is biased. Why should a single process model show bias?

There are at least three general conceptualisations of the truth bias. It may arise from an error in the system, just as a dual process account would also claim, it may reflect a useful aid to help navigate the social world, or it may be a way of making informed, accurate decisions based on an integration of context-relevant knowledge with currently available information. The three possibilities I am suggesting are considered in more detail in the next section, and will be empirically tested in Chapters 4 and 5. I will conclude that the last of these conceptualisations best explains the nature of the truth bias. I will show this by demonstrating that the other two accounts fail to stand up to empirical testing and that the adaptive use of context-relevant knowledge has a causal effect on the truth bias.

In short, I will argue for a single-process account of the lie-truth judgment and suggest that the truth bias emerges from an integration of context-relevant knowledge and behavioural information from the speaker in order to form a satisfactorily accurate judgment. This conclusion contributes to both the decision-making and lie detection literatures. For the former, I argue that context-relevant information is brought to bear on real-world complex decisions, where cues are available not only from the immediate environment, but also from prior knowledge and from an understanding of the social environment. This thesis contributes to the lie detection literature by borrowing from decision-making research a new theoretical perspective, the adaptive decision-maker account (Payne, Bettman & Johnson, 1993), considered in more detail in Chapter 2.

Conceptualising the Truth Bias

There is a substantial body of evidence demonstrating the existence and the pervasiveness of the truth bias (Bond & DePaulo, 2006, 2008; DePaulo, Charlton, Cooper, Lindsay & Muhlenbruck, 1997; Levine et al., 1999; Vrij, 2008; Zuckerman et al., 1979; Zuckerman, DePaulo & Rosenthal, 1981). At least three conceptions of the truth bias have been made: (i) as an error in the system that directly causes a decrease in accuracy (see Vrij, Granhag & Porter, 2010), (ii) as a result of socialisation practices that, as a by-product, causes a decrease in accuracy (O'Sullivan, 2003), and (iii) as an adaptive strategy that makes use of contextual knowledge and the available behaviour of the speaker to make as good a judgment as possible given the constraints of the limited cognitive resources and the limited information in the world (Payne et al., 1993; Simon, 1990).

First, it has been thought of by some as an aberration in the system (see Vrij, 2008; Vrij, Granhag & Porter, 2010), in much the same way as a bias in a weighing scale can be thought to be an unwanted but systematic problem with an otherwise useful device. Raters are thought to have an overgeneralised bias towards believing because of the lack of appropriate knowledge (Fiedler & Walka, 1993), for example, or because raters neglect some of the available information (Ask & Granhag, 2007; see also Vrij, 2008) and tend to anchor to that information in spite of contradictory evidence (Elaad, 2000; Fan, Wagner & Manstead, 1995; Zuckerman, Koestner, Colella & Alton, 1984; see also Gilbert, 1991). An important part of improving accuracy is to overcome the truth bias, from this perspective.

In contrast, others have considered the truth bias not as a fault or defect, but as a useful aid in a social world. Accusing others of being a liar is a socially aggressive act (O'Sullivan, 2003). And it may come with potentially aggressive repercussions, from a short-lived argument to long-term breakdown of relationships (Bell & DePaulo, 1996; Clark & Lemay, 2010; Cole, 2001; DePaulo & Bell, 1996; Guthrie & Kunkel, 2013; Miller, Mongeau & Sleight, 1986). It may be in a rater's interest to sacrifice accurate lie detection to maintain social cohesion. The truth bias is not an inherent part of the system but is rather an 'optional extra' that we bring to the task.

A third way has also been suggested: the truth bias is useful, but need not decrease accuracy. In fact, it is present precisely because raters are trying to improve their accuracy. People tell the truth far more often than they lie (DePaulo, Kashy, Kirkendol, Wyer & Epstein, 1996), so it would make sense (in terms of increasing accuracy) to hedge on the side of believing others (O'Sullivan, Ekman & Friesen, 1988). In this sense, the truth bias is more of an adaptive and flexible response that emerges from a strategy of incorporating prior knowledge into the decision. Police interviewers, for example, show a bias towards *disbelieving* others (Kassin, Meissner & Norwick, 2005), which may reflect their tendency to expect their interviewees to lie to them (Kassin, 2005; Masip, Alonso, Garrido & Antón, 2005). Such a lie bias is also seen when the speaker is believed to be generally untrustworthy (see DePaulo & DePaulo, 1989).

Thus the truth bias can be thought of in at least three different ways: either as an aberration in the system, a willing sacrifice of accuracy to abide by social conventions, or an adaptive strategy that incorporates prior knowledge and expectations. This thesis will consider which of these three conceptualisations can best describe the truth bias. It will be argued that the first two are untenable. The truth bias appears to be a flexible response in that biased responding adapts to situational information and thereby improves performance. How these different conceptualisations will be addressed is discussed further below and in greater detail in Chapters 2 and 3.

Whether the truth bias is considered an aberration, a social convention, or an adaptive strategy addresses what Marr (1982) called the computational level of analysis. It gives an insight into the function of the system. But answering this question tells us nothing of the algorithmic level – how the system carries out its operations to perform these functions.

Process Accounts of the Truth Bias

Given the applicability of lie detection research, it has been primarily geared toward finding new methods to improve accuracy. As a result, there have been few attempts to understand the underlying cognitive processes involved (see Miller & Stiff, 1993; Reinhard & Sporer, 2010; Vrij & Granhag, 2012).

Attempts to understand the processes are only beginning to be considered. A number of researchers have begun to show support for a dual process theory of the truth bias (Gilbert, 1991; Gilbert, Krull & Malone, 1990; Masip et al., 2006; Masip, Garrido, et al., 2009; Masip, Garrido & Herrero, 2010; Reinhard & Sporer, 2008, 2010). Dual process theory has had success in explaining a wide range of social-cognitive phenomena, and has had some preliminary support in being able to account for lie-truth judgments (Brocas & Carrillo, in press; Chaiken, Liberman & Eagly, 1989; Chaiken & Trope, 1999; Evans, 2008; Evans & Stanovich, 2013; Fiori, 2009; Gawronski, Sherman & Trope, in press; Gilbert, 1991; Granhag, 2006; Masip et al., 2006; Masip, Garrido, et al., 2009; Masip et al., 2010; Petty & Wegener, 1999; Reinhard, 2010; Reinhard & Sporer, 2008, 2010; Smith & DeCoster, 2000; Ulatowska, 2013). I will empirically test the claims made by one prominent dual process theory, the heuristic-analytical model (HAM), and determine whether it can account for the truth bias.

The HAM makes a distinction between a fast, effortless but error-prone heuristic processing stream and a slower, more evaluative but effortful analytic processing stream. It is thought that the truth bias arises from the more effortless heuristic processing route. Three general classes of HAMs have been identified (Evans, 2007). Two of these, the default-interventionist and the parallel-competition models, propose that with a short amount of processing time the heuristic route will be used, but that with longer durations a more analytical processing style can be used. Accounts of the truth bias have taken this tack (Gilbert, 1991; Gilbert et al., 1990; Gilbert, Tafarodi & Malone, 1993; Masip et al., 2006; Masip, Garrido, et al., 2009; Masip et al., 2010). The third class of models, pre-emptive conflict resolution models, have been applied to questions concerning accuracy rather than bias (Reinhard & Sporer, 2008, 2010). These models propose either heuristic or analytical processing is chosen at the outset of the judgment, rather than switching between them after time has elapsed. The types of information in the environment may necessitate the selection of analytical processing, although internal motivations can also be used to make the initial selection of heuristic or analytical processing, according to this class of model.

These three classes of model are discussed in greater detail in Chapter 2. The distinction between them is important for the current purposes because it allows for testable predictions to be generated from what is a broad and general theoretical framework (see Evans & Stanovich, 2013). A number of challenges will be brought to bear on the HAM, briefly outlined below and in more detail in Chapter 2. I will show that there is no support for a dual-process HAM in accounting for the bias. Instead, my research suggests the truth bias can better be accounted for by a single process that makes use of context-dependent information and available cues in the environment in order to form relatively simple judgments in an uncertain environment. I will argue we are single-process adaptive decision makers (Beach & Mitchell, 1978; Gigerenzer & Selten, 2001; Gigerenzer, Todd & The ABC Research Group, 1999; Payne et al., 1993; Platzer & Bröder, 2012; Simon, 1990) that in an information-limited world with a resource-limited cognitive capacity make satisfactorily accurate judgments.

In the next section I address how I intend to examine the underlying processes and the conceptual nature of the truth bias.

Testing the Truth Bias: Process and Purpose

Here I shall give an overview of the structure of this thesis. I will first give an account of the general argument I wish to make, and will then follow with an outline of how each study contributes to the narrative.

This thesis will begin by considering whether the HAM can account for the truth bias. It has been shown that during the early moments of consideration raters are truth biased, but that over time the truth bias is attenuated (Gilbert et al., 1990; Gilbert et al., 1993; Masip et al., 2006; Masip, Garrido, et al., 2009). Over time the heuristic processing stream is thought to be interrupted by an analytical process, thereby reducing the bias (Masip et al., 2006, 2009, 2010).

If the HAM can explain the presence (and absence) of the bias as a switch in processing modes, it must stand up to a number of challenges. First, it must be shown that the decrease in the truth bias is not simply the result of the speaker changing in their behaviour over time.

Second, because two of the three classes of HAMs make claims based on processing durations, the decline in the truth bias must be predicted by the amount of processing time available. A switch between the heuristic and analytical processing streams is thought to occur after a given amount of processing time, at least under two of the three general classes of HAM.

This brings us to the third challenge. If processing time cannot predict the decline in bias, then only one class of the HAM remains. This class of models predicts a selection of heuristic or analytical processing from the outset. Speech requires attention to the narrative of the story and is considered to necessitate analytical processing (Forrest & Feldman, 2000; Gilbert & Krull, 1988; Reinhard & Sporer, 2008, Study 2; Stiff et al., 1989). Visual cues are considered easier to process (Forrest & Feldman, 2000; Reinhard & Sporer, 2008, Study 3; Stiff et al., 1989) and so should evoke heuristic processing.

If none of the above challenges can be met, an alternative account that makes no claim to multiple processes may better explain the phenomenon. I explore the possibility that raters bring contextually relevant knowledge to bear on the decision given the difficulty with and uncertainty in making a lie-truth judgment. I will show that raters do so precisely because they are uncertain but nonetheless have to make some decision. They can rely on relatively simple rules that the current context allows. If multiple statements are available from the speaker, for example, comparisons between statements can be made, and a consistency heuristic (‘inconsistent statements are deceptive statements': Granhag & Strömwall, 1999) can be used. In the absence of multiple statements, raters can make use of other speaker-specific information such as how likely a given person is to lie, how alike the speaker is to the rater, and so on. Raters use this information when they must make a judgment but are unsure and so have to guess. The central theme of this thesis is that raters do not flip a mental coin – they make use of the available information, in the form of the speaker's behaviour and from context-specific knowledge, to make an *informed* guess.

Testing the claims

A more detailed account of how I intend to empirically test the preceding argument can be found in Chapters 2 and 3. Here I will give a brief overview as to how I intend to empirically address the challenges I have posed above.

The first three experiments of Chapter 4 will begin by determining whether the three classes of the HAM can account for a decline in the truth bias as a speaker delivers a number of statements (Masip et al., 2006; Masip, Garrido, et al., 2009). The first experiment sought to determine whether the truth bias could be explained as a simple behavioural shift on part of the speaker. Shifting around the order of the statements should not affect the declining truth bias phenomenon because it is assumed to be independent of the behaviours being judged. This was supported. However, the degree of truth bias was not dependent on the amount of processing time available, failing to support two of the three classes of HAMs.

Experiment 2 addressed the third remaining class of models by manipulating the types of information available to the rater. This has previously been shown to determine whether heuristic or analytical processing is selected. If this account were to explain the truth bias, it would have to be the case that for the speaker's first statement the rater made use of the visual information, thought to be processed heuristically, but over time changes to verbal information, thought to be processed analytically. However, regardless of the type of information presented (only audio or only video), the decline in the truth bias was observed, failing to support the final class of HAMs. In addition, the amount of processing time available was again unable to predict the degree of truth bias, replicating Experiment 1. Experiment 3 explicitly manipulated the amount of processing time available, rather than relying on observational analyses. There was no effect of viewing time on the degree of bias.

Having established that the HAM fails to meet the challenges issued, an alternative account was sought that does not rely on a distinction between processing modes: the use of a consistency heuristic (Experiment 4). It was shown that consistency was not only able to account for the decline in bias across my research, but also for the decline in bias found in the original research (Masip, Garrido, et al., 2009). Thus the decline in bias can be explained by the use of a relatively simple heuristic.

It is not possible to rule out time-based HAM accounts altogether, however. It may be argued that before the end of the first statement raters may have shifted from a heuristic to analytical process. Having found no support for a HAM account at the coarse time scale of minutes, in Chapter 5 a greater temporal resolution is explored, at the scale of seconds (Experiment 5). Others have reported an early truth bias during the act of comprehension (Gilbert et al., 1990; Gilbert et al., 1993). In Experiment 5 I replicate this effect. Based on the findings of Experiment 4, again it is considered whether raters are making use of a relatively simple heuristic. Although consistency between statements is not possible, raters can make use of other readily accessible and context-relevant information such as their prior experience with the world to make 'the best guess' in the absence of any other differentiating behaviour from the speaker. Experiment 5 finds that when forced into judgment, raters show a truth bias during the early moments of processing. The early truth bias depends on how likely they believe the speaker is lying or telling the truth (Experiment 6).

The HAM failed to meet the challenges issued: the truth bias can better be accounted for in ways that do not rely on a dual process distinction. The HAM distinction could explain *why* there was a truth bias. Dropping this model leaves us without an explanation of why the truth bias exists. Does it arise from the single-process cognitive structure in some way, does it reflect socialisation practices, or rather does the bias merely emerge as an adaptive response to the available information under uncertainty? The findings of Experiment 4, showing that it arises from a consistency heuristic when it is an available option, and the findings of Experiments 5 and 6, showing raters make use of base rate information when it is necessary to guess, suggest an adaptive role of the truth bias. For example, when raters expected speakers to lie, there was an early bias towards *disbelieving* the speaker.

Experiment 7 particularly highlights the adaptive nature of the truth bias, albeit from a post-hoc position. The purpose of this study was to extend the findings of Experiment 5, which found an early truth bias during comprehension. Because a truth bias is observed after comprehension and at the point of judgment response, the aim was to show that the truth bias at the point of the final judgment was also a result of making the best guess. By removing the need to make a judgment, it was thought that the bias would be attenuated. However, the exact opposite was found: when raters no longer needed to make a judgment they showed a greater truth bias. Further exploration of the data soon revealed that those forced to make a judgment used their uncertainty in a rather strategic way. Uncertainty *after the statement and behaviours had been comprehended* is different in nature to uncertainty because of a lack of information. The former is uncertainty *in light of* the available evidence. This in itself can be a useful guide. Raters are better at spotting truths than lies, a phenomenon known as the veracity effect (Levine et al., 1999). Perhaps unsurprisingly then they are more confident when making truth than when making lie ratings (Anderson, DePaulo & Ansfield, 2002; DePaulo, 1992; DePaulo et al., 1997; DePaulo et al., 2003; Anderson, 1999, cited by DePaulo & Morris, 2004; Hartwig & Bond, 2011; see also Levine et al., 1999). If raters were still unsure after having evaluated the behaviour, this likely suggests the statement is not a truth. Thus the 'best guess' under uncertainty in this instance is not to guess truth, but rather to guess lie. This can be seen clearly in the behavioural data: raters forced to make a judgment rate speakers as truthful just as often as those not forced into judgment, but they do rate speakers as lying more often. Those not forced into judgment would explicitly indicate their uncertainty, whereas those forced into judgment were more likely to err on the side of disbelieving. The best guess is not always to guess the speaker is telling the truth: in some situations it is more adaptive to guess the speaker is lying.

The findings thus far appear to align with an adaptive decision-making perspective. Yet it may be that a social account can better explain the truth bias. Consider that raters may not be guessing 'truth' because they are relying on their prior knowledge of the world but because of their understanding of the social rules: it is rude to call someone a liar. The aim of Chapter 6 was to determine whether implied social presence alone was sufficient to result in a truth bias.

If the truth bias results from treating others favourably because of our understanding of the social situation, removing the social component of the situation should remove the bias. To test this, all audio and most of the video cues were removed from the stimulus, leaving only a wire frame outlining the movement of the speakers. Raters were led to believe these wire frames were either social or non-social beings. Specifically, they were either told the speakers were participants in a police interview situation or were told that the wire frame outputs were the result of computer modelling of a collection of thousands of behaviours and that each video produced a set of behaviours typical of deception or truth-telling. Believing that the videos were of other human speakers was not sufficient to invoke a truth bias (Experiment 9). While no support was found in favour of such a role for social norms, it is suggested that social information can be utilised as a simple cue to help guide the judgment, as shown in Experiment 10.

Chapter 7 and 8 conclude with a discussion of the work presented and its implications for the HAM and the adaptive decision-maker. I argue context-relevant information is important for making informed judgments. In natural environments outside of the laboratory context-relevant information can be available in the immediate environment but we can also draw on our prior experiences and our understanding of the social situation in help guide the decision.

Chapter 2: Theoretical Accounts of the Truth Bias

Chapter 1 gave a brief overview of the aims and intentions of this thesis: to explore decision-making under uncertainty in a socially oriented task. In that chapter I showed how heuristic-analytic models (HAMs) have had some success in accounting for biased (Gilbert, 1991; Masip et al., 2006; Masip, Garrido, et al., 2009) and (in)accurate (Reinhard & Sporer, 2008, 2010; see also Fiori, 2009; Ulatowska, 2013) lie-truth judgments. In this chapter I will consider in detail the predictions made by the dual-process HAM and flesh out how these predictions will be tested in the experiments described in Chapters 4 and 5. I will also consider in greater depth the position taken by the single-process adaptive decision-maker account, which is thought to use relatively simple heuristics to make informed decisions under uncertainty.

The HAM is particularly prominent in the persuasion literature. It can explain when a person is more or less likely to be persuaded (i.e. when heuristically processing: Chaiken, 1980; Chaiken et al., 1989; Petty & Cacioppo, 1986; Petty & Wegener, 1999). The distinction between a fast and frugal process versus a slow and effortful process has long been considered in decision-making research (Evans, 2008; Gilovich et al., 2002; Kahneman & Tversky, 1973). Surprisingly, it is only recently that HAMs been applied to understanding both biased and accurate lie-truth judgments (Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993; Masip et al., 2006; Masip, Garrido, et al., 2009; Masip et al., 2010; Reinhard & Sporer, 2008, 2010), perhaps because of the lack of process-oriented research in the area (see Lane & Vieira, 2012; Miller & Stiff, 1993). This thesis will critically evaluate the utility of HAMs in explaining the truth bias by generating a set of testable predictions that any HAM must meet. These tests are made on two different but related phenomena, examined separately in Chapters 4 and 5.

This chapter will give greater attention to the distinctions made between three classes of HAMs and highlight the predictions made by each. A set of challenges will be issued to test these models. These challenges are (1) the truth bias must at least in part arise from the cognitive operations of the rater, (2) the amount of processing time available must be able to predict the shift in biased responding (as predicted by two of the three classes of HAMs), and (3) if it cannot, then the types of information available in the environment should guide an early selection of either heuristic or analytical processing (as predicted by the third class of HAMs). It will be shown that the HAM fails to meet these challenges. A single process adaptive decision-maker account is a viable alternative account, one that makes use of simple heuristics.

Note that a distinction is being made between heuristic processing, a claim of the HAM, and the use of heuristics, employed by both the heuristic-analytic and the adaptive decision-maker accounts. There is potential for confusion between heuristics and heuristic processing. This can be seen in the literature where the terms are sometimes used interchangeably. The distinction is made explicit in the following section.

Definitions: Heuristics, Heuristic-Analytical Processing and The Adaptive Decision Maker

Unfortunately, the term *heuristic* has been used in markedly different ways – and has even been used interchangeably with the term *heuristic processing* (Gigerenzer & Gaissmaier, 2011). In fact, the term heuristic has been so widely used that its meaning has become vague and blurred (Shah & Oppenheimer, 2008). As a result, heuristics are considered consistent with almost all research findings, but they make few testable predictions (Gigerenzer & Gaissmaier, 2011). To overcome this, I will give a working definition of heuristics as tools built up from past experience in similar contexts. Both HAMs and the adaptive processing account make use of these experience-based heuristics, but they can also make use of other frugal strategies. The two accounts differ insofar as they make different assumptions about the number of processing routes and about the utility of heuristics as either error-prone or beneficial.

Heuristic Tools

Heuristics are defined as simple generalised rules of thumb built up through individual, evolutionary and/or social experience, and are used in forming a decision, particularly under uncertain conditions (Gigerenzer & Gaissmaier, 2011; Gilovich et al., 2002; Hutchinson & Gigerenzer, 2005; Simon, 1990; Tversky & Kahneman, 1974). The identifying component of this definition is a history of interaction with the world (Gigerenzer & Gaissmaier, 2011; Kruglanski, 1989). The availability heuristic, for example, can be used to estimate how frequently an event or class of events occur by making use of how often instances of the class can be brought to mind from memory (Tversky & Kahneman, 1974). That instances of plane crashes from news media reports can be brought to mind from easier than can instances of car crashes may lead to the (erroneous) inference that plane crashes are the more frequent. Heuristics through experience give rise to simple rules such as 'librarians are quiet people' and 'older people are more likely to have heart attacks'. Heuristics like these can be thought of as a set of tools that can aid decision-making insofar as they offer a quick and ready solution to a given problem. But because they necessarily oversimplify the state of the world by generalising from past experience, heuristics can sometimes lead to erroneous inferences (see Gilovich et al., 2002).

Note that where HAMs would propose a duality in processing styles, one of which is a heuristic processing stream, no such distinction is being made in the definition of a heuristic. Heuristics are the generalised rules of thumb that can be used by whatever cognitive processes are thought to be in use.

Heuristic and Analytical Processing

Heuristics are entirely independent of the heuristic-analytic processing distinction. *Heuristic processing* is one mode through which information is evaluated, characterised primarily by its ease and speed of processing. The process can make use of the heuristic tools described above, but it can also (or instead) use other means to arrive at a quick and relatively effortless judgment (Chaiken et al., 1989; Chaiken & Trope, 1999), such as one's own moral code (Haidt, 2001), the types of information available in the environment (Chaiken, 1980; Chen & Chaiken, 1999; Forrest & Feldman, 2000; Gilbert & Krull, 1988; Petty & Wegener, 1999; Reinhard & Sporer, 2008; Stiff et al., 1989), whether one is guided by accuracy concerns or other goals like social group acceptance or self-affirmation (Bohner, Moskowitz & Chaiken, 1995; Chen, Shechter & Chaiken, 1996; Wyer & Frey, 1983; see also Martin &

Hewstone, 2003), or other accessible attitudes relevant to the judgment (Houston & Fazio, 1989; see also Howard-Pitney, Borgida & Omoto, 1986). That is, readily accessible information, whether from the environment, past experience or self-reflection, can be used to make relatively effortless decisions.

As well as using readily accessible information, heuristic processing could take into account *all* the available information in the environment and from past experience, much as might be expected of a more analytical and deliberative form of processing, but could exert less effort in coming to a judgment (Ajzen & Sexton, 1999; Chaiken, 1980; Griffin, Neuwirth, Geise & Dunwoody, 2002). Heuristic processing, then, is a relatively effortless way of gathering and manipulating information to arrive at a judgment. It can do so by making use of heuristic rules, but could also do so by other means, such as more shallow processing of the available information (Ajzen & Sexton, 1999; Chaiken, 1980) or relying on a select set of readily accessible cues in the environment (Chaiken et al., 1989; Chaiken & Trope, 1999).

Analytical processing is an alternative mode of processing under the HAM. It is more deliberative and evaluative than heuristic processing, meaning it processes data in more depth and with more effort, but it too can make use of heuristics (Bohner et al., 1995; Chaiken et al., 1989; Chaiken & Maheswaran, 1994; Chaiken & Trope, 1999; Petty, Cacioppo & Goldman, 1981; Wood, Kallgren & Preisler, 1985; Worth & Mackie, 1987). Heuristics, although coarse, are informed by prior experience with the world. For that reason they can be informative for more reasoned decisions (Chaiken et al., 1989). Thus whether heuristic rules are used or not does not define whether a process is said to be a heuristic process or an analytic process. Rather, it is the depth and the evaluative nature of the processing effort that distinguishes heuristic processing from analytical processing (Ajzen & Sexton, 1999; Griffin et al., 2002; see also Kahlor, Dunwoody, Griffin, Neuwirth & Giese, 2003). However, it is difficult to define precisely what is evaluative and what is not evaluative, or to draw a clear demarcation between 'deep' and 'shallow', for example. This difficulty in clearly distinguishing between the modes of processing continues to be a source of debate (Evans & Stanovich, 2013; Keren, 2013; Keren & Schul, 2009; Kruglanski, Pierro, Mannetti, Erb & Chun, 2007; Osman, 2004; Thompson, 2013), and is discussed in more detail in Chapter 7. For now, deeper processing is taken to mean more cognitive effort is expended in evaluating the information.

To summarise so far, heuristics are rules of thumb that are informed by past experience. They are just one of the set of tools that can be used to help judgment formation under conditions of uncertainty. They can be used whether using heuristic or analytical processing, although their effects tend to be less influential on the latter mode. Heuristic processing is a more shallow and less evaluative deliberation that makes use of readily accessible information (Ajzen & Sexton, 1999; Chaiken, 1980): sometimes, the most readily accessible information is a heuristic. This *heuristic tool* versus *processing mode* distinction should be borne in mind.

The Adaptive Decision-Maker

The adaptive decision maker is defined as one who makes functional links between the limited information that can be obtained, from both the environment and past experience (Gigerenzer & Brighton, 2009; Gigerenzer & Goldstein, 1996; Gigerenzer, Martignon, Hoffrage, Rieskamp & Czerlinski, 2008; Payne et al., 1993; Simon, 1990), with the limited cognitive resources available to the decision maker (Chase, Hertwig & Gigerenzer, 1998; Gigerenzer & Goldstein, 1996; Shah &

Oppenheimer, 2008; Todd & Gigerenzer, 2005), such that, in general, decision outcomes will be successful (e.g., Marewski, Gaissmaier & Gigerenzer, 2010). This definition closely aligns with that of Payne et al. (1993). In the current context of lie detection, successful decisions are thought to be accurate ones.

The adaptive decision maker has been considered in learning and decision-making tasks, where raters can make use of multiple cues to inform their decision (e.g., Bröder, 2000, 2003; Bröder & Eichler, 2006; Gigerenzer & Goldstein, 1996, 1999; Gigerenzer & Selten, 2001; Goldstein & Gigerenzer, 2008; Platzer & Bröder, 2012). These studies find raters limit their attention to a select number of cues in order to form a satisfactorily accurate judgment. As yet, the adaptive decision maker has not been applied to more socially oriented tasks like the lie detection judgment, which too are multiple-cue decision tasks.

To make successful decisions with ambiguous and limited information, cognitive attention is guided towards only a limited set of cues. Some research suggests the selection of limited cues is a deliberative decision rather than an automatic tendency (Bröder, 2003; Platzer & Bröder, 2012). The selection may be determined by the demands of the task (Payne et al., 1993; Rieskamp & Hoffrage, 1999), but when undetermined, the more accessible and salient information may be used (Newell & Shanks, 2003; Platzer & Bröder, 2012).

The adaptive decision maker account (Beach & Mitchell, 1978; Gigerenzer & Selten, 2001; Payne et al., 1993; Platzer & Bröder, 2012; Simon, 1990), however, brings together aspects of both heuristics and heuristic processing. Briefly, it shares with the HAM that frugal strategies most effectively employ limited cognitive resources (Gigerenzer & Goldstein, 1996; Gigerenzer et al., 2008) by using simplified strategies in complex, information-limited worlds (Gigerenzer et al., 1999; Todd &

Gigerenzer, 2005). Unlike the HAM distinction, though, the adaptive thinker need not possess two processing streams, and the use of simplified rules is considered to be functional and effective rather than maladaptive and error-prone.

In summary, as with heuristic processing, heuristics could be one of the tools found in the adaptive decision-maker's toolbox, although other means of simplifying the decision process are also available (Gigerenzer & Goldstein, 1999). Unlike the dual-process HAMs though, the adaptive decision-maker account that I propose is a single process theory, making no claim to a second, qualitatively different style of processing (see Gigerenzer & Gaissmaier, 2011 for a similar perspective that makes no link between the use simple rules with a distinct unconscious or associative processing mode). In addition, contrary to the view of the effortless and error-prone heuristic process in dual-process accounts, the adaptive decision-maker simplifies the uncertain environment in order to *improve* decision-making by making use of relatively simple rules and reducing cognitive attention to only a select portion of the information set available (Chase et al., 1998; Czerlinski, Gigerenzer & Goldstein, 1999; Gigerenzer, 2007; Gigerenzer & Gaissmaier, 2011; Gigerenzer et al., 1999; Goldstein & Gigerenzer, 2008; Simon, 1990). Where heuristic processing is considered to make *oversimplifications* of the available evidence, the adaptive decision-maker is considered to employ generally *successful shortcuts* in an information-limited world.

Summary of Definitions

I have drawn a distinction between heuristics, heuristic processing, and the adaptive decision-maker. Heuristics are generalised rules of thumb built up from experience, such as the availability heuristic. Heuristic processing makes use heuristics and other available information to arrive at a relatively effortless and potentially erroneous decision, whilst analytical processing is a more deliberative form of information processing. The adaptive decision-maker account shares properties of both heuristics and heuristic processing: it makes use of simplified rules that, in an uncertain world, are generally successful. It can make use of whatever information is readily accessible, akin to a heuristic process, but differs inasmuch as (a) there is no claim to a second more effortful processing mode and (b) makes use of simplified rules because they can increase judgmental accuracy in an information-limited environment, rather than because of a cognitive miserliness.

This thesis will conclude that the heuristic-analytical dual-process model cannot account for the phenomena that, on the surface, appear to exhibit the patterns of a dual processing structure. Heuristics informed by prior experience appear to be used in some, but not all circumstances: readily accessible cues in the environment can also be used for fast and frugal decision-making. The adaptive decision-making account proposes a set of generally successful strategies that rely on limited amounts of information, either from prior experience or in the immediate environment, under conditions of uncertainty. I will argue the adaptive decision-making account best explains the findings of this thesis, and that even uncertainty itself can be used in an adaptive way to aid decision-making. First though, I will begin from the theoretical framework that has been applied to understanding the truth bias: the heuristic-analytic model.

A Dual Process: The Heuristic-Analytic Model (HAM) in Detail

Dual process theories, as the name suggests, have in common the proposal that there are two modes or two stages of information processing that can lead to different judgment outcomes. A prominent model from the persuasion literature is the heuristic-analytic model (HAM, e.g. Chaiken et al., 1989; Chaiken & Trope, 1999; Evans, 2007). The HAM shares with other dual-process theories a claim to two processing streams. Heuristic processing results in fast, intuitive judgments (De Neys, 2006; Evans, 2007; Sloman, 1996) and has the advantage of consuming little cognitive resources (Chaiken et al., 1989). Unfortunately it is prone to neglect the informative evidence available (Chaiken et al., 1989). But we are not confined to heuristic processing: a second analytical mode of processing is available for more analytical deliberation. Analytic processing is slower, requiring greater effort, cognitive resources and cognitive ability (e.g. Chaiken et al., 1989; Petty & Cacioppo, 1986).

Because these processing streams are different in the way they select and process information, an important concern for dual process theories is how to address the potential for conflicting outcomes of the two systems of thought. In a review of HAMs, Evans (2007) identified three main classes that dual process theories could be categorised into based on the way in which they deal with potential conflict: (1) pre-emptive conflict resolution, (2) parallel-competition, and (3) default-interventionist models. This taxonomy is adopted here for two reasons. First, they identify a small set of essential features that make up HAMs in general. The predictions of each of these class of models can be empirically tested and allow for a more general conclusion regarding the applicability of HAMs to be drawn. Should no support be found for these three classes of models, then inferences regarding HAMs in general can be made (although see Evans & Stanovich, 2013, for a claim that HAMs are not directly Popperian falsifiable but rather require a prolonged attack at the Kuhnian core of the theory). Secondly, and relatedly, they are sufficiently broad as to encapsulate the many theoretical approaches and psychological disciplines reviewed by Evans (2007). A lack of support for any one class of model will have implications for the many theories that adopt the principles of that class of model. With this in mind, the current thesis will examine each of these classes of model in an attempt to distinguish which, if any, class of model may best account for the truth bias.

HAMs have seen some success in their application to understanding lie-truth judgments. Whilst questions about lie detection *accuracy* have been met with pre-emptive competitive HAMs (Reinhard, 2010; Reinhard & Sporer, 2008, 2010; Stiff et al., 1989), the issue of the *truth bias* has seen application of the default-interventionist class of model, if only implicitly (Fan et al., 1995; Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993; Masip et al., 2006; Masip, Garrido, et al., 2009; Zuckerman, Fischer, Osmun, Winkler & Wolfson, 1987; Zuckerman, Koestner, et al., 1984). In this section I will explore all three classes of model and consider whether there are alternative explanations of the reported phenomena that better account for the truth bias. Two accounts of the truth bias rest on the claim that heuristic processing is followed by analytical processing, a prediction consistent with both the parallel-competition and default-interventionist models. As such, focus will be given to these models, although pre-emptive conflict resolution models will also be considered.

Three classes of HAMs

(1) *Pre-emptive conflict resolution* models address potentially conflicting judgment outcomes by avoiding the possibility in the first place (Evans, 2007). Either heuristic or analytical processing is selected from the outset. This selection may be constrained by the information available in the environment. Analytical processing is engaged where more effortful processing will be required (Chaiken, 1980; Evans, 2007; Evans, Newstead & Byrne, 1993). Speech, which requires the listener to comprehend and reconstruct the narrative, requires greater cognitive resources (Forrest & Feldman, 2000; Gilbert & Krull, 1988; Reinhard & Sporer, 2008, Study 2; Stiff et al., 1989) than do visual behaviours, which have been shown to be easier to process (Forrest & Feldman, 2000; Reinhard & Sporer, 2008, Study 3; Stiff et al., 1989). When only verbal information is present, analytical processing is engaged (Chaiken, 1980; Chen & Chaiken, 1999; Petty & Wegener, 1999). When only visual information is present, the heuristic system is chosen (Reinhard & Sporer, 2008, Study 3; Stiff et al., 1989).

The selection of heuristic or analytical processing need not be driven by the environment. For example, analytical processing has been shown to be the more likely processing route when motivation or task involvement is high (Chaiken, 1980; Chaiken et al., 1989; Chaiken & Maheswaran, 1994; Chaiken & Trope, 1999; Chen & Chaiken, 1999; Forrest & Feldman, 2000; Petty & Cacioppo, 1979, 1986; Reinhard, 2010; Reinhard & Sporer, 2010). In the absence of environmental cues that can deterministically lead to a selection of a processing route, internal factors can motivate the choice of a process. Whether determined by the external world or selected by internal motivations, pre-emptive conflict resolution models circumvent the potential for conflict between the two processing modes at the earliest stage of forming a judgment, before it has begun.

(2) *Parallel-competition models* propose heuristic and analytical processing routes run in unison (Evans, 2007). The mediator of conflict in this model is time. As noted previously, analytical processing is slower and more deliberative. The heuristic processing stream is quicker because it is less analytical and more global in its evaluation. The heuristic processing stream will thus finish before the slower analytical route. Where a quick decision is desired or required, the heuristic processing stream is *de facto* selected because the analytical processing stream has yet to produce an output. However, where more time is available for judgment the analytic process is favoured. Sloman's (1996, 2002) associative-rule-based model captures this distinction well.

(3) *Default-interventionist models* share this time property. This class of models also proposes heuristic processes will be the basis of judgment if a quick judgment is made. The model differs from the parallel competition model insofar as the former propose the heuristic system is the default mode of processing and that after some given period of time it is interrupted by analytical processing (Evans, 2006, 2007; Stanovich & West, 2000), rather than the two processing modes running simultaneously. Both default-interventionist and parallel-competition models receive support in various domains from research showing a response bias when quick judgments are made but a reduced bias when additional processing time is available, suggesting a primacy of heuristic processing (De Neys, 2006; Evans & Curtis-Holmes, 2005; Gilbert et al., 1990; Gilbert et al., 1993; Masip, Alonso, Garrido & Herrero, 2009; Masip et al., 2006; Roberts & Newton, 2001; Schroyens, Schaeken & Handley, 2003; Verschueren, Schaeken & d'Ydewalle, 2005).

For example, in one study, Gilbert and colleagues presented learners with a set of nonsense words and a supposed definition of that word (Gilbert et al., 1990, Study 1). The definition was indicated to be either true or false. If participants were given little time to encode the veracity of a definition, they more often believed the definition was accurate and true. On those occasions where they had longer to learn whether the definition was true or false, they were less likely to be biased. The authors interpreted this as evidence of an initial default processing stage that automatically encodes information as true in the first instance, but with sufficient time a second stage of processing was thought to be intervening and revising their belief (Gilbert, 1991; Gilbert et al., 1990).

In summary, three classes of HAMs can be identified. Pre-emptive conflict resolution models select either heuristic or analytical processing at the outset. Parallel-competition models propose the two run in tandem and that whether the analytical processing stream has had enough time to arrive at a judgment determines whether the heuristic or analytical processing outcome is used. Finally, the default-interventionist account similarly proposes time constraints on the use of heuristic and analytical processing, but claims the heuristic process is a default mode of encoding and that at later stages analytical processing can intervene upon. This distinction offers clear predictions that the HAM makes and allows us to empirically ask whether the more general HAM framework is a useful account of the truth bias.

Challenges for a Heuristic Processing Account of the Truth Bias

Having given a detailed account of the three classes of HAMs, I will now begin to issue a set of challenges that they must meet in order to be able to account for the truth bias.

The multi-response interview: A declining truth bias

In a set of experiments, Masip and colleagues (2006, 2009, 2010) interviewed speakers about a mock crime they had just watched on video. The interview consisted of three questions that asked about the actions of each of the three characters in the mock crime. The three responses came together to form a single statement, in much the same way that police interviewees may be asked multiple questions about different aspects of a crime which would be taken as a single statement about the course of events. Speakers were instructed to either lie throughout their statement (i.e. across all three responses) or to tell the truth throughout. In one study, participants were presented with the recordings of each speaker's three responses (Masip et al., 2006). They were to make a lie-truth judgment at the end of the statement and to indicate whether they come to their decision by the first, second or third response. Raters were more inclined to believe the speaker if they reported making their decision earlier. In another study, raters gave a judgment after each response, thereby providing three judgments for each speaker (Masip et al., 2009) rather than a single response as before. Replicating the effect, raters showed a more marked truth bias after the initial response but that by the second and third response the truth bias had declined. It was also found raters became more accurate over time, as one might expect if raters shifted from an error-prone heuristic process to a more analytical mode of thought.

These studies are novel inasmuch as they accept lie-truth judgments can change over time as more information becomes available. The reported phenomenon, a decreasing truth bias, gives a window onto the changing nature of the truth bias and exposes the conditions under which a truth bias is more or less likely to be present, in this instance early or late in the judgment process, respectively. Thus it offers a promising place from which to start exploring the underlying cause of the bias.

Although the authors acknowledged that their 'results are open to alternative interpretations' (Masip et al., 2010, p. 591), they strongly favoured a HAM interpretation. I will suggest and test two alternative accounts for these findings: the behavioural account and the single-cue account. These two explanations along with the HAM are addressed below.

The behavioural account of the truth bias. Labelling excessive truth responding as a bias implies it is a simple rule-based tendency on part of the rater that is prone to error. Given this interpretation, perhaps it is unsurprising it has come to be thought of as an erroneous outcome of heuristic processing. Yet it might be the case that the truth 'bias' is not a cognitive bias at all, but instead an unbiased and valid inference guided by the behaviours available to the rater. Surprisingly, such an explanation has been little considered (although see Fan et al., 1995; Zuckerman et al., 1987; Zuckerman, Koestner, et al., 1984, for some discussion of this possibility). Both liars and truth-tellers deliver believable statements in an attempt to appear (rightly or wrongly) honest. That is, truth tellers expect to be believed by others (Gilovich, Savitsky & Medvec, 1998), whilst liars manipulate their behaviour to appear as though they are speaking the truth (Buller & Burgoon, 1996; DePaulo et al., 2003). Whether lying or telling the truth, we might expect the speaker to exhibit or attempt to exhibit behaviours that suggest they are telling the truth. Raters' judgments may reflect an astute consideration of the speaker's apparently genuine behaviour (Fan et al., 1995; Zuckerman et al., 1987; Zuckerman, Koestner, et al., 1984), resulting in a high degree of truth responding. Lie responding may similarly be guided by the presented behaviour: when the speaker's behaviour is unexpected and violates norms, they are more likely to judge the statement as deceptive (Bond et al., 1992; Burgoon & Walther, 1990). And as these behaviours change over time, so too do the ratings of that behaviour (Buller & Burgoon, 1996; Chung & Fink, 2008; Weld & Danzig, 1940; see also Saykaly, Talwar, Lindsay, Bala & Kang, 2013).

Consider that the truth bias has been shown to decline over the course of the statement and become more accurate in their judgments (Masip et al., 2006, 2009, 2010). This may be an indication of, for example, liars becoming more nervous over time, or leaking cues to their deception in other ways that are picked up by the rater. Either more diagnostic behaviour or merely more deceptive behaviours displayed later in the statement would result in a decline in the observed truth bias. Truth-tellers on the other hand, generally confident in the belief that 'the truth will out' (Gilovich et al., 1998; Hartwig, Granhag & Strömwall, 2007; Kassin, 2005; Moston, Stephenson & Williamson, 1992), may not become more nervous over time and will not leak cues to deceit any more than would be expected by chance. In this way, as time progresses liars and truth-tellers will become more distinct from one another (Granhag & Strömwall, 2002). Truth-tellers' behaviours would be perceived as honest across their statement whereas liars would become increasingly unconvincing, mirroring the improved accuracy and reduced bias found in prior studies (Masip et al., 2006, 2009).

If a heuristic-analytic account is to receive any support, it is necessary to show that the decline in the truth bias over time results from the listener's developing considerations *independently* of the specific behaviours that are being rated. That is, there must be evidence of an independent cognitive component that operates irrespective of the behaviour presented. In Chapter 4 I will begin by exploring whether there exists just such an independent cognitive effect.

The truth bias as a decline across time, not across ratings. Prior studies have shown a decline in the truth bias between the first, second, and third response of the speaker (Masip et al., 2006, 2009). While this offers a nominal sense of progression, it is the case that some speakers would have provided particularly long responses whereas others were rather short. The difference in the duration of statements is attributable to the fact that the researchers collected spontaneously generated lies and truths. Because analytic processing intervenes after a given duration of processing *time* rather than after a given number of ratings (see Evans, 2007), it is important that the differences in viewing durations, which will differ depending on the length of the speaker's statement, are not glossed over by considering them to have been sampled after the same amount of processing time. To support a HAM account, it is important to show the truth bias decreases as the cumulative duration of the speaker's statement increases. This may seem like a trivial distinction. However, the act of making a rating is itself an influence on what judgment is reached (Granhag & Strömwall, 2000b; Hogarth & Einhorn, 1992). This is considered further in the next section.

In their work, Masip and colleagues (2006, 2009) considered the truth bias to be attributable to the default-interventionist class of HAMs. In order to more thoroughly test the applicability of the HAM to the truth bias, all three classes will be considered here. Both the default-interventionist and the parallel-competition models make the same claim regarding processing duration: the analytical process will have a greater influence on the judgment outcome with longer processing durations. These accounts can thus be tested simultaneously.

The third class of models supposes a selection between heuristic or analytical processing before the judgment process proceeds. That there is a demonstrable change in the judgment outcome over time has in the past been taken as evidence that weakens the pre-emptive conflict resolution model precisely because it makes no temporal predictions (Evans, 2007). These models instead predict that internal motivations or the cognitive effort required to process particular types of information in the environment allow for an early selection of processing mode. Because the environment can be more readily manipulated and has been argued to deterministically result in heuristic or analytical processing, this approach will be taken in Chapter 4 to test the claims of this class of models.

The single-cue account of the truth bias. In the preceding section I briefly alluded to research showing how the act of making multiple judgments can itself be a causal force in the decision process (Granhag & Strömwall, 2000b, 2001a; Hogarth & Einhorn, 1992). In order to understand this effect, it is worth considering how judgments are made when multiple responses are available.

The act of rating after each new response provides not only additional processing time but also new affordances for judgment. When two or more responses have been provided, it is possible to make comparisons between them. When the opportunity for comparison arises, raters have been shown to make use of consistency in the statement more so than other available cues such as the amount of detail they provided and even the plausibility of the statement (Granhag & Strömwall, 1999, 2000b, 2001b; Strömwall & Granhag, 2005, 2007; see also Strömwall, Granhag & Jonsson, 2003).

Unfortunately (for the rater), consistency appears to have low diagnostic utility in distinguishing adults' lies from truths (Granhag & Strömwall, 1999, 2001b, 2002; Strömwall et al., 2003; although see Strömwall & Granhag, 2007, for evidence of consistency as a predictive cue of children's lies). Surprisingly, there is no clear agreement between raters whether a given speaker appears consistent or inconsistent (Granhag & Strömwall, 1999), although there is general agreement that inconsistency is a cue to deception (Granhag & Strömwall, 1999, 2000a, 2001a). Because raters perceive inconsistencies even when they are not present, there could be a shift towards rating others as deceptive when the opportunity for comparison arises, i.e. after having viewed the second response. Having established inconsistency after the second response, there may be no additional effect of inconsistency by the third response: the lack of consistency may carry over from the second to the third response. In the prior studies it was found that the decline in truth bias was concentrated between the first and second response; there was no further decline by the third response.

This account raises two empirical challenges for the HAM. First, it must be shown that the amount of processing time, independent of the number of judgments made, can predict the decline in truth bias. Second, the decline in truth bias must not be attributable to the *consistency heuristic* (Granhag & Strömwall, 1999), but instead reflect a shift from a heuristic process to an analytical process.

Because heuristic processes can make use of heuristics, it may be questioned whether this is really a challenge to the HAM. Strictly speaking, by itself it is not a strong challenge to the HAM. However, the phenomenon of a reduced truth bias has been taken as evidence in favour of a HAM. If the phenomenon can be accounted for without reference to the HAM, in unison with other findings that directly challenges it, this finding further undermines the dual process account by suggesting a new interpretation of the phenomenon. These challenges will form the remainder of Chapter 4.

Summary: Challenges for the HAM account. This thesis sets out to test the applicability of HAMs to the truth bias. One recently identified phenomenon aligns with the default-interventionist and parallel-competition class of HAMs: raters show a truth bias initially that, over the course of the statement, declines (Masip et al., 2006, 2009, 2010). In this section I presented some challenges that the HAM account must overcome. First, it must be shown that the truth bias is not solely a direct reflection of the speakers' behaviours. If the truth bias is a product of a more error-prone processing route then the rater is the cause of his or her own bias, regardless of the behaviours of the speaker. In this sense, it must be shown there is an 'independent cognitive component' to the truth bias and that it is not a veridical reflection of the behaviour the speaker portrays. It is important to note that focus is given to cognitive dual processing accounts over, say, emotion accounts (e.g., Fiori, 2009). The nature of this first challenge is to address whether the cause of the truth bias resides at least in part with the rater. Discussion of an 'independent cognitive phenomenon' serves as a useful shorthand through this thesis. However, it should be noted that if the evidence suggests this challenge can be met, this does not then entail evidence in favour of a cognitive mechanism over an emotional mechanism, for example.

Second, it must be shown that the declining bias is predicted by processing time and not simply by the number of judgments made. Because statements can be long or short, the number of ratings acts only as a proxy to viewing time, so it cannot directly address the predictions of HAMs. This second challenge assumes either a default-interventionist account, as the authors suggested (Masip et al., 2006, 2009), or a parallel-competition account. If neither of these accounts receive empirical support, a third challenge must be met instead if HAMs are to have any explanatory power: it must be shown, in accordance with a pre-emptive conflict resolution model, that the nature of the information available to the rater will influence whether raters are truth biased or not. Verbal information, requiring greater cognitive processing, should reflect an analytical processing style and result in a reduced truth bias. Each of these challenges is addressed in the first empirical chapter, Chapter 4.

If these challenges cannot be met, an alternative account of the phenomenon may shed light on the causal factors behind the truth bias. One possibility discussed in this section arises from the affordances arising from the task. By receiving multiple responses from the same speaker about characters involved in the same crime scene, it is possible to start making comparisons. Perceived inconsistencies in statement may result in a reduced truth responding. Such a finding would suggest judgmental heuristics are to account for the truth bias. This would be consistent with an adaptive decision-making account, where the decision outcome results from attention to a limited set of cues in the environment.

Hope for the HAM: Processing at a finer time scale

The truth bias has been discussed above with reference to the declining truth bias across multiple judgments of a speaker's statement. I will argue with the use of this phenomenon that a HAM cannot explain the truth bias. There are two major drawbacks of using this phenomenon to test HAMs. First, although it may be possible to explain the decline in the truth bias as a shift from heuristic to analytical processing, it is not immediately clear *why* heuristic processing shows this primacy. Second, although a failure to meet any of the challenges proposed earlier certainly weakens the HAM's position, it is not quite possible to rule it out entirely. This is because of the unspecified time scales at which the heuristic and analytical processes are said to operate on the truth bias. Whilst the pre-emptive conflict resolution model is clear in this respect – the choice of processing mode occurs at the start of the judgment – the remaining two classes of model suggest no specific time frames at which the analytical process will be favoured over the heuristic process. Although HAMs may fail to account for the declining truth bias at the coarse time scale of minutes, accounts that predict more fine-grained smaller time scales of seconds are still tenable.

A Spinozan mind: Believing in order to comprehend. One prominent account of this sort stems from the work of the Dutch philosopher Baruch de Spinoza (1677/1982). Drawing on his work, Gilbert and colleagues (Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993) proposed the truth bias is a necessary outcome given the structure of the mind. The 'Spinozan mind' hypothesis claims that in order to comprehend a statement, it is necessary to first believe the statement is true. This view continues to impact philosophy (Burge, 1993; Millikan, 1987) and other domains such as hypothetical reasoning (Fitzsimons & Shiv, 2001; Moore, Neal, Fitzsimons & Shiv, 2012), text comprehension (Hasson, Simmons & Todorov, 2005; Prentice, Gerrig & Bailis, 1997; Schul, Mayo & Burnstein, 2004), persuasion (Green & Brock, 2000;

Prentice et al., 1997; Sperber et al., 2010) religious belief (Pennycook, Cheyne, Seli, Koehler & Fugelsang, 2012), credibility assessment (Colwell et al., 2012; Levine et al., 1999; Millar & Millar, 1997), and other areas of research (Chen & Blanchard-Fields, 2000; Knowles & Condon, 1999; Skurnik, Yoon, Park & Schwarz, 2005).

According to the Spinozan mind hypothesis, it is only after the initial acceptance that we can consider rejecting the idea we initially assented to believe. Thus comprehension is considered a two-step process where the 'unbelieving' stage follows automatic acceptance in time (Gilbert et al., 1990). As such, it is a prototypical example of a default-interventionist model where a default belief stage is superseded by a more deliberative consideration, making it a suitable candidate for testing HAMs. Because comprehension is online, which is to say that we comprehend each new piece of information as and when it becomes available (e.g. Heuttig, Rommers & Meyer, 2011; Spivey, Grosjean & Knoblich, 2005; Spivey, 2007), the effects of initial belief due to comprehension and the subsequent revision of information are proposed to take place within the initial few seconds of receiving information (Gilbert et al., 1990). This allows us to tackle the issue that may be levied at the multiple-response interview paradigm discussed above, where it could be claimed the time scale of minutes is too coarse to detect the shift from heuristic to analytical processing.

In favour of the Spinozan view that belief comes prior to disbelief is research showing that negated sentences, e.g. 'the eagle was not in the sky', are slower to process than affirmed sentences, e.g. 'the eagle was in the sky' (Carpenter & Just, 1975; Clark & Chase, 1972; Clark & Clark, 1977; Donders, 1969; Mayo, Schul & Burnstein, 2004; Trabasso, Rollins & Shaughnessy, 1971; Zwaan, Stanfield & Yaxley, 2002). It has been suggested that information is encoded initially as true and that negation requires additional lexical ‘not’ tags to identify the statement as false (Clark & Chase, 1972; Clark & Clark, 1977; Gilbert, 1991; see also Johnson-Laird, 1983; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998), just as the Spinozan mind hypothesis predicts.

However, recent research has started to show that it is the nature of the task, not the structure of the mind, that results in faster processing for affirmed statements. Stating that ‘the eagle was not in the sky’ leaves open various possibilities as to where the eagle was: in the nest, on the floor, on a dinner plate, and so on. Theories of comprehension such as mental model theory (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991) or those that rely on embodiment theory (Zwaan et al., 2002) suggest a representational ‘image’ of the scene is constructed online. Thus where there is ambiguity, such as is the case in underspecified negated statements like ‘the eagle was not in the sky’, mental reconstruction of the scene requires a consideration of all the possible alternatives, which in turn will appear to slow the comprehension process. It has been shown that when the negated statements are not underspecified, they are processed equally as quickly as affirmed statements (Anderson, Huette, Matlock & Spivey, 2009, 2010; Glenberg, Robertson, Jansen & Johnson-Glenberg, 1999; Wegner, Coulton & Wenzlaff, 1985; see also Deutsch, Kordst-Freudinger, Gawronski & Strack, 2009; Schul, Mayo & Burnstein, 2004. Also see Dodd & Bradshaw, 1980, and Skowronski & Carlston, 1989, for a similar account of impression formation). People seem to adapt to the structure of the task.

The automatic encoding of information should also impact the way new information is sought, according to the Spinozan claim (Gilbert, 1991). It is suggested people test hypotheses about the world by seeking evidence that confirms their beliefs and hypotheses about the world. For example, given the hypothesis ‘Bob is introverted', it is claimed people seek out information that confirms this view rather than attempt to test it by looking for disconfirming evidence. Indeed, the confirmation bias is a well-documented phenomenon (see Snyder & Campbell, 1980; Snyder & Swann, 1978). Yet the hypotheses in these studies are of the form that would be employed if one already knew the hypothesis to be true, e.g. 'Bob is introverted'. It would be normative to ask questions that confirm the statement because they can give us more detail about what is already known, such as *how* introverted Bob is (Gilbert, 1991; Higgins & Bargh, 1987; see also Trope & Bassok, 1982).

The disconfirmatory hypothesis, that Bob is an extravert, is usually only implicit and never directly given to participants. In their review of confirmation biases, Higgins and Bargh (1987) note that a number of studies show no preference for confirming evidence when the disconfirmatory hypothesis is explicitly presented to participants ('Bob is extroverted'). As with the processing of negation, the preference for confirming one's beliefs only exists when the confirmatory hypothesis has a single construal (e.g. 'I am introverted') and the alternative disconfirmatory hypothesis is underspecified, meaning there is a range of potential disconfirmatory hypotheses that could be explored (e.g. 'Bob is extraverted in some situations', 'Bob is extraverted around certain people', 'Bob is only a little extraverted').

Upon examination, it becomes clear that prior research does not immediately lend itself towards favouring the Spinozan view, as Gilbert et al. (1993) have also noted. Gilbert and colleagues (1990, 1993) more directly address the empirical predictions of their Spinozan mind account. Because believing is said to be an initial default state, the claim can be tested from two angles. First, if belief precedes more evaluative processing, interrupting the judgment process early versus allowing additional time should result in a tendency towards believing, as indeed they showed

(Gilbert et al., 1990, Study 1). Second, if it is a default state, then mere comprehension should be sufficient to result in belief (Gilbert et al., 1990, Study 3). Similarly, adding a secondary task should increase the cognitive load and lead to reliance on the default state (Gilbert et al., 1990, Study 2; Gilbert et al., 1993). Thus they concluded there is strong support for a Spinozan view of the mind.

A Cartesian mind: A single-process alternative. This view can be contrasted with Descartes' (1641/1993). The 'Cartesian Mind' is able to comprehend information independently of assessing its veracity. Under this view, there would instead exist an initial period of indecision and a subsequent evaluation (Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993). That is, where a Spinozan mind posits an early acceptance, the Cartesian mind posits an initial uncertainty. Because there is only a single judgmental process, it is considered a one-stage model in contrast to the Spinozan two-stage dual process model (Gilbert, 1991; Gilbert et al., 1990). Gilbert and colleagues rejected the Cartesian mind because, in their view, it is not capable of accounting for an early truth bias.

The rejection rests on the assumption that uncertainty will be exhibited by respondents as being equally likely to believe as to disbelieving, in line with the logic behind the two-alternative forced choice. I will call this the naïve Cartesian model because it assumes that under uncertainty raters respond randomly. However, decision strategies can employed by a decision-maker under uncertainty, and can be used to offer an 'informed guess'. The Cartesian view can be modified to incorporate these strategies, which will be called the informed Cartesian model. Under this view, comprehension begins with a period of uncertainty. Knowledge and past experiences can bias initial uncertainty *towards* believing a statement, before a firm decision is made whether to believe or doubt. If the decision-maker is forced into judgment despite not having made one yet, they can rely on their past knowledge of the world.

The availability heuristic that 'people usually tell the truth' may be relied upon to help make this forced decision (see O'Sullivan et al., 1988; Tversky & Kahneman, 1974). Because heuristics are the result of past experience with the world, they can be informative: studies show that the frequency at which truths are encountered far outweighs the frequency of lies we experience (Caspi & Gorsky, 2006; DePaulo & Kashy, 1998; DePaulo et al., 1996; George & Robb, 2008; Hancock, Thom-Santelli & Ritchie, 2004; see also Cole, 2001; van Swol, Malhotra & Braun, 2012). The informed Cartesian mind would be expected to show a truth bias during the early moments of comprehension because it makes use of its prior knowledge. Both an informed Cartesian and a Spinozan mind can account for the findings.

The informed Cartesian claim implies that the bias might not be constantly set to affirm automatically, but rather is a preferential bias that is dependent on experiential or situational factors. Pre-existing choice preferences can come from experience and expectations (Beukeboom, Finkenauer & Wigboldus, 2010; Deutsch et al., 2009; Hanks, Mazurek, Kiani, Hopp & Shadlen, 2011; Hasson et al., 2005; Schroeder, Richter & Hoever, 2008; Schul, 1993; Schul et al., 2004; van Ravenzwaaij, Mulder, Tuerlinckx & Wagenmakers, 2012), and regulations constrained by the situation ('innocent until proven guilty': Pennington & Hastie, 1991), which have been shown to play an important role in comprehension (Glenberg & Robertson, 1999, 2000).

There is some evidence that this preference can be modified, which, importantly, should not be possible of an automatic Spinozan bias towards acceptance. State suspicion shifts the bias towards doubting (Deutsch et al., 2009;

Kassin et al., 2005; Kim & Levine, 2011; Masip et al., 2005; Moore et al., 2012; Schul et al., 2004; see also Bond, Malloy, Arias, Nunn & Thompson, 2005), negated statements can be processed faster than affirming statements when the uncertainty inherent in the task is removed (Anderson et al., 2009, 2010; Glenberg et al., 1999; Hasson et al., 2005; Wegner et al., 1985; see also Beukeboom et al., 2010; Fraundorf, Benjamin & Watson, 2013; Glenberg & Robertson, 2000), and forewarning increases the tendency towards disbelieving (Allyn & Festinger, 1961; Benoit, 1998; Hovland & Weiss, 1951; Kiesler & Kiesler, 1964; McGuire, 1964; Wood & Quinn, 2003). Thus under different conditions people can actually appear *anti*-Spinozan.

To recap, the Spinozan mind hypothesis is a form of default-interventionist dual process model that makes predictions about the use of heuristic versus analytical processing on the time scale of seconds. Whilst a HAM account cannot explain the observed decrease in the truth bias found at the more coarse time scale of minutes, it may have success on this time scale. The advantages of using the Spinozan model are (1) it has received empirical support in other domains, (2) it continues to influence research across a variety of areas, and (3) if support is not found for a Spinozan model an obvious alternative candidate is readily available: the informed Cartesian rater that relies on past experience. Chapter 5 tests the predictions of each of these models.

The findings of Chapter 5 will show how raters are not dual-process Spinozans, but are informed single-process Cartesians, making adaptive decisions under uncertainty. The concept of uncertainty plays a central role, and a distinction between internal and external uncertainty is made. Internal uncertainty reflects the inability to decide whether to believe or disbelieve because the information gathered from the environment is not sufficiently diagnostic to push us towards one judgment or the other.

Internal uncertainty, it is argued, can be used adaptively. Raters are typically less confidence when listening to lies than truths and are less confident in making lie judgments compared to making truth judgments (Anderson et al., 2002; DePaulo, 1992; DePaulo et al., 1997; DePaulo et al., 2003; Anderson, 1999, cited by DePaulo & Morris, 2004; Hartwig & Bond, 2011; see also Levine et al., 1999). Consider also that raters are better at spotting truths than spotting lies (Levine et al., 1999). Similarly, processing affirmed statements, e.g. 'the eagle is in the sky', are easier and take less processing time than negated statements, e.g. 'the eagle is not in the sky' (Anderson et al., 2010; Carpenter & Just, 1975; Clark & Chase, 1972; Glenberg et al., 1999; Johnson-Laird, 1983; Trabasso et al., 1971). In both cases, internal uncertainty can be used as an indicator that the statement is likely to be false.

External uncertainty results from a lack of information available in the environment on which to make a decision. Under external uncertainty, raters can come to rely on their prior knowledge of similar situations, or more general information about the current context. An availability heuristic such as 'people usually tell the truth' (Grice, 1975; O'Sullivan et al., 1988), and an understanding that communication of new information needs to be *true* information if it is to be useful (see Fiedler, Armbruster, Nickel, Walther & Asbeck, 1996; Grice, 1975; Swann, Guiliano & Wegner, 1982), can be used to guide the decision, and may result in a bias towards believing. The use of an availability heuristic under external uncertainty is explored in Experiment 6 and considered further in the next chapter.

Conclusion: Chapter 2

To account for how we come to make decisions when situated in the real world, this thesis makes use of a social paradigm in which the information in the environment (the behaviours of the speaker) is ambiguous, creating a degree of uncertainty in the decision. Despite the inherent ambiguity, there is a systematic tendency towards believing others. This thesis takes as its starting point a theoretical framework that has been applied to the field of lie detection: the heuristic-analytic model.

Four challenges were brought to bear against the HAM account: (1) the truth bias must have some cognitive component independent of the presented behaviour, (2) cumulative processing time must be able to predict the decline in the truth bias, (3) If processing time is unable to account for the decline in bias, then it must be shown that heuristic or analytical processing is chosen from the outset, and (4) If the HAM cannot account for the truth bias at a coarse temporal resolution, it should be able to do so at a more fine-grained. It will be shown that there is little evidence to favour a dual process account. Instead, evidence is offered to suggest raters are adaptive decision makers in the face of uncertainty and can make use of information from the environment, from prior experience or even make use of the uncertainty itself to aid decision-making. In the next chapter, the types of information used by an adaptive decision-maker are considered.

Chapter 3: Context-Dependent Information Use

I have so far considered a prominent process theory that has been applied to understanding lie detection: the heuristic-analytic model. I have argued that a HAM cannot explain the truth bias and that it may be better explained with an adaptive decision-maker account. This account claims that we make flexible use of a limited amount of context-relevant information in the environment and from past experience to aid making a decision under uncertainty. For example, participants in a learning task may discover that although all the cues are diagnostic, some are more diagnostic than others, and so will limit their attention to the more diagnostic ones in making their decision.

This chapter considers what information may be used in the absence of information in the environment. An alternative to the adaptive decision-maker account will also be considered: the truth bias may simply reflect social norms of not accusing others of being a liar.

The current chapter will argue that in the absence of behavioural cues, raters can make use of more general information such as base rate knowledge and information about the social world. However, it will be argued that while raters may make use of social information in an adaptive fashion, it is not the case that the truth bias reflects learnt socialisation practices of politeness.

The Availability Heuristic: Base Rate Knowledge

In exploring cognitive HAM accounts, whether at the longer time scales found across speaker's testimonies (the multiple-response interview) or at more fine-grained time scales during the process of comprehension (the Spinozan mind hypothesis), the evidence appears to suggest an overweighting of a single cue compared to the use of other available cues in order to make a judgment. It is suggested that when there is no information available in the environment, raters can make use of their prior knowledge and experience with similar situations. In this section the role of the availability heuristic considered in preceding chapter – i.e. 'most people tell the truth' – will be more directly considered.

'Most people tell the truth'

Believing others may be an accurate reflection of the real world. The average person lies on a daily basis, but the number of deceptive interactions is far outweighed by the number of honest interactions we experience (Caspi & Gorsky, 2006; Clark & Clark, 1977; DePaulo et al., 1996; Fan et al., 1995; O'Sullivan, 2003; O'Sullivan et al., 1988). One possible reason for this asymmetry is that a high rate of honesty may be necessary to successfully communicate and to be understood (Echterhoff, Higgins & Levine, 2009; Grice, 1975). On those relatively rare occasions when people do lie the quality of the conversation is reported to be less pleasant and intimate (DePaulo et al., 1996; Miller et al., 1986), which may also serve to discourage deception. Whatever the reason, the true base rate of honesty is likely far higher than the equal split of lies and truths presented in the laboratory (see DePaulo et al., 1996; O'Sullivan et al., 1988). As such, the ‘people rarely lie’ heuristic is likely a fair representation of the world.

An informed decision-maker should factor this knowledge into their judgment (Nisbett & Ross, 1980); after all, labelling most of our interactions as deceptive when there are likely to be few of them will result in high inaccuracy. Outside of the laboratory, this heuristic can serve as a useful aid to forming accurate judgments: anticipating that others will be honest is likely adaptive when an individual has encountered more honest than deceptive interactions in their daily lives. This information, it is predicted, is most likely to have an influence when there is little information available in the environment, i.e. under external uncertainty.

There is some evidence suggesting the expected rate of honesty influences the truth bias. Participants made aware that they might be deceived rate speakers as more deceptive compared to naïve raters (Stiff et al., 1992; Toris & DePaulo, 1985), and as suspicion increases, statements are less likely to be rated as truthful (e.g. Bond et al., 1992; Levine et al., 2000; Masip, Alonso, et al., 2009; McCornack & Levine, 1990). That is, ratings appear to reflect expectations of deceit. It is also worth noting that police investigators, who have a lie bias (Meissner & Kassin, 2002), expect their interviewees to lie to them (Kassin et al., 2005; Masip et al., 2005; Moston et al., 1992). Lay raters similarly show a bias towards disbelieving others when those others are believed to be sales people (DePaulo & DePaulo, 1989), which the authors explained might reflect their perceptions of how honest these individuals typically are. Their expectations built up from past experience may be the cause of their biased responding (see Kassin, Goldstein & Savitsky, 2003). The use of simple rules such as these built up from past experience characterises a heuristic.

However, what is not clear is how suspicion comes to influence the judgment. Suspicion of deception may increase motivation to seek out the liars compared to those who are not suspicious. Or it may be that suspicion is used as a heuristic cue to the base rate: warned of the possibility of deception, raters may come to rely less on their general belief in others' honesty and instead come to use a lie heuristic such as 'most of *these* people do not tell the truth'. That is, it is unclear whether suspicion increases motivation to detect liars or updates the perceived base rate.

The adaptive decision-maker offers a specific time-course prediction in making use of base-rate information: when little information is available from the environment, during the earlier moments of comprehension, the provided base rate information should be used because it is the only available evidence about the current speaker's veracity. Over time, as information becomes available, the base rate may come to have less of an impact on the judgment.

That the truth bias is a direct reflection of the base rate is by no means a new suggestion (DePaulo & Rosenthal, 1979; O'Sullivan et al., 1988). To my knowledge, though, there has been no direct test of the claim. To more directly assess whether the rate of honesty is being used as a relatively simple cue to an individual's honesty, Experiment 6 will explicitly manipulate beliefs about the base rate of honesty, tracking over the course of the decision how the base rate information is used.

Socialisation Practices

Thus far consideration has been given to behavioural and cognitive accounts. But believing and disbelieving others is an inherently social task. From an early age we are taught that lying is morally wrong (Backbier, Hoogstraten & Terwogt-

Kouwenhoven, 1997; The Global Deception Team, 2006; see also Boon & McLeod, 2001; Guthrie & Kunkel, 2013), but it is encouraged when it can grease the wheels of social life (DePaulo et al., 1996; Lewis, 1993; Roggensack & Sillars, 2013; Vrij, 2008; Vrij, Granhag & Porter, 2010). The final empirical chapter considers an alternative to the adaptive decision-maker position that has been advocated thus far. I will consider how the social world influences how we come to believe, and whether it can account for the truth bias. This section will begin by examining the accusatory reluctance account (DePaulo & Rosenthal, 1979; Ekman, 1992; Miller et al., 1986; O'Sullivan, 2003; O'Sullivan et al., 1988; Vrij, 2008; Vrij, Granhag & Porter, 2010), which proposes raters are willingly naïve in order to avoid an aggressive social act. To my knowledge, it has never been empirically tested. After considering this account, preliminary evidence will be presented suggesting that use of even a single social cue, similar to the overuse of a single cognitive cue considered earlier, may result in biased decision making.

Accusatory reluctance

Directly confronting someone we think is lying is a socially aggressive act. Accusing someone of lying challenges the claim, their integrity, and the relationship. In the short term, the other person could respond with equal social aggression, whether an argument and a defence of their claim or an attempt to socially distance themselves from the accuser. In the long term, the relationship may be harmed or even break down (Bell & DePaulo, 1996; Clark & Lemay, 2010; Cole, 2001; DePaulo & Bell, 1996; Guthrie & Kunkel, 2013; Miller et al., 1986; Roggensack & Sillars, 2013; although see Aune, Metts & Hubbard, 1998, for coping strategies upon discovery of deception). There are clear negatives, then, both short and long term (see Sagarin, Rhoads & Cialdini, 1998).

The positives are not so evident. In cases of serious deception, such as the denial of a theft, the obvious benefit is the return of the stolen item. But in less serious cases there may be no clear benefit, and as discussed there are likely to be negative outcomes associated with it (Vrij, 2008). For example, having given an unwelcomed gift that is accepted with thanks and feigned pleasure, spotting the lie means having to accept your chosen gift was not well thought out, the thanks undeserved and the monetary value lost. In this situation, we typically feign pleasure in receiving an unwanted gift (The Observer, OM Magazine, 11 January 2004, p. 12, cited by Vrij, 2008). In a more consequential scenario, people use lying as a means of avoiding destructive arguments in an attempt to maintain relationships (Bell & DePaulo, 1996; Cole, 2001; DePaulo & Bell, 1996; DePaulo et al., 1996; Guthrie & Kunkel, 2013). Looking for the truth may damage the relationship. This wont to believe others and to not actively seek out the truth has been called by Vrij (2008) the Ostrich effect (see also DePaulo et al., 2003; Ekman, 1992; Vrij, Granhag & Porter, 2010).

Thus there are tangible social benefits that accompany a willingness to sacrifice lie detection accuracy in favour of believing others. As such, a number of authors have noted these social practices may result in a bias towards taking others' claims at face value (DePaulo & Rosenthal, 1979; Ekman, 1992; Miller et al., 1986; O'Sullivan, 2003; O'Sullivan et al., 1988; Vrij, 2008; Vrij, Granhag & Porter, 2010). The accusatory reluctance account (most explicitly outlined by O'Sullivan, 2003) proposes listeners are biased towards believing others because the act of accusing another is an aggressive social act that breaks with societal rules and norms (Vrij, Granhag &

Porter, 2010), and in turn can result in equally aggressive social repercussions for the accuser.

Although the accusatory reluctance has been proposed as a possible explanation when discussing the truth bias, the account has never been empirically tested. However, there is some empirical evidence that offers support for the position. A reluctance to label others as liars presupposes that we can detect deception at rates that are higher than have been typically reported. After all, while the experimenter may consider high accuracy the normative standard, participants may feel the pragmatics of social interaction, social perception and attribution to be more important forces in how they form their judgments (Fiske, 1992). Ekman and colleagues (Ekman, Friesen, O'Sullivan & Scherer, 1980; O'Sullivan, Ekman, Friesen & Scherer, 1985) showed that raters attend to different sources of information available from the speaker depending on whether they were listening to lies or truths, suggesting they are making a distinction between these types of communications that are not explicitly reported. Similarly, Hurd and Noller (1988) found raters are more likely to consider the possibility of deception in their verbal reports whilst listening to a lie than whilst listening to a truthful account. Also, there is less trust placed in those who lie even when listeners do not explicitly acknowledge they have been lied to (Sagarin et al., 1998).

The accusatory reluctance account also suggests raters should feel less comfortable when listening to a lie than when listening to a truth because of the apprehension associated with accusing another person. Indeed, research shows there is greater discomfort and less confidence when listening to lies than when listening to truths (Anderson et al., 2002; DePaulo et al., 1997; DePaulo & Kashy, 1998; see also Toris & DePaulo, 1985).

If the truth bias is predicated on our understanding of the social environment, we may also anticipate that as the nature of the social relationship augments, so too does the willingness to accuse another. Partners are less likely to give negative feedback to those with whom they are in close relationships compared to more distant relational partners (DePaulo & Bell, 1996; DePaulo & Kashy, 1998; Lemay & Clark, 2008; Uysal & Oner-Ozkan, 2007; see also Boon & McLeod, 2001), and when it is given it is often subtle (Clark & Lemay, 2010; Metts, 1989; Swann, Stein-Seroussi & McNulty, 1992).

Thus there is evidence to suggest raters may be reluctant to accuse others of deception, and that this reluctance could stem from socialisation practices. That is, a case can be built for the view that raters are unwilling to call out others on their lies because of the rules and norms that surround our perceptions of others as social agents. In Chapter 6, I directly test this claim.

Social relatedness as an adaptive strategy

Information about the social world can be used in a less socially oriented fashion. Rather than bringing a willing naïveté to the task, social cues, like others discussed already, could be used as a heuristic to making a judgment. There may be no such apprehension to label social agents liars, but knowing that a speaker is psychologically close (e.g. a relational partner) rather than psychologically distant (e.g. a stranger) may be sufficient for causing a truth bias.

It has been argued that availability heuristics such as ‘people like me tell the truth’ and ‘people close to me are trustworthy’ are used when rating close relational partners (McCornack & Parks, 1986; O'Sullivan, 2003; Stiff et al., 1992; Wickham, 2013). For example, there is a generalised tendency to trust partners when we feel closer to them (O'Sullivan, 2003; Stiff et al., 1992; Wickham, 2013). As the relationship becomes closer and as the length of the relationship increases, there is an even greater bias towards believing them (Argyle & Henderson, 1984; McCornack & Levine, 1990; McCornack & Parks, 1986; Stiff et al., 1992; Stiff et al., 1989; van Swol et al., 2012; see also Boyes & Fletcher, 2007) and a greater tendency to trust and be more cooperative with them (Gaertner & Dovidio, 2000; Voci, 2006). This *social heuristic* account claims social relatedness information can be used to *aid* decision-making under uncertainty. In contrast, the accusatory reluctance account claims our understanding of the social world actively guides us towards *inaccurate* decisions in favour of abiding by social norms.

Much research has explored the biasing effects of social relatedness. The minimal social grouping paradigm (Tajfel, 1970; Tajfel, Billig, Bundy & Flament, 1971) isolates the effects of group belonging to study how even the most minimal degree of relatedness to others can influence our behaviour. Because assignment to groups is random, the only link between participants within a given group and the only difference between their group (known as the ingroup) and the other group (known as the outgroup) is the fact that they have been arbitrarily categorised. This paradigm has allowed researchers to examine the effects of social categorisation *per se*, independent of other features of grouping that tend to co-occur naturally such as attitude similarity between ingroup compared to outgroup members. As such, the paradigm allows researchers to explore how social relatedness information is used independently of other factors.

This research consistently shows that we treat members of our own group more favourably than outgroup members (Billig & Tajfel, 1973; Brewer, 1979, 1999; Brewer & Silver, 1978; Gaertner & Dovidio, 2000; Moy & Ng, 2006; Otten &

Mummendey, 2000; Tajfel et al., 1971; Voci, 2006), in line with the social heuristic account (McCornack & Parks, 1986; Stiff et al., 1992). In one such experiment, Tajfel et al. (1971, Study 2) presented participants with reproductions of abstract paintings by the modern artists Paul Klee and Wassily Kandinsky. They were led to believe they would be grouped based on their preference for one of the two artists, but in reality were randomly assigned to either the Klee preference or Kandinsky preference group. Having been assigned a group, participants distributed monetary reward to others. Despite there being no cost associated with rewarding both ingroup and outgroup members equally nor any benefit for rewarding the two groups unequally, there was nonetheless a tendency to favour the ingroup (see also Brewer & Silver, 1978). Similar findings have been shown when the grouping was known to be based on the flip of a coin; that is, when participants were aware the grouping was in fact random rather than based on preferences (Billig & Tajfel, 1973). The mere act of being grouped with others is sufficient to cause changes in interpersonal behaviour.

These findings suggest a preferential treatment of ingroup members rather than (or in addition to) a derogation of outgroup members. In three reviews of intergroup bias, the same conclusion has been reached: there is an asymmetry between ingroup preference and outgroup derision, such that members are more likely to show preferential treatment to ingroup members than they are to actively derogate outgroup members (Brewer, 1979, 1999; Otten & Mummendey, 2000).

As noted, the advantage of the minimal group paradigm is that the perceived relationship between two individuals can be studied independent of other factors that tend to co-occur with grouping, such as similar preferences on the grouping dimension (Allen & Wilder, 1979; Diehl, 1989; Tajfel et al., 1971), an understanding of social rules built between them (Vrij, 2008), or even the ease of processing the speaker's voice which may cause a sense of fluency (see Alter & Oppenheimer, 2009; Begg, Anas & Farinacci, 1992; Whittlesea, 1993). To date, studies showing a positive relationship between social relatedness and the truth bias have confounded these factors. This gives rise to two related concerns. First, the truth bias may be entirely independent of social relatedness, and instead attributable to, say, processing fluency. That is, we may give people we know the benefit of the doubt because we have experience with their accent and tonality, creating a sense of processing fluency that has been shown to be related to being truth biased (see Reber & Schwarz, 1999; Schwarz, Bless, et al., 1991; Unkelbach, Bayer, Alves, Koch & Stahl, 2010). Second, even if there is a causal connection between relatedness and the truth bias, it is not clear whether it is a sufficient condition by itself or whether relatedness combines with these other factors that in turn leads to the emergence of a truth bias. That is, it is unclear whether relatedness is a sufficient condition to cause a truth bias.

In Chapter 6 I will utilise the minimal social grouping paradigm as employed by Tajfel and colleagues (1971) to test the social relatedness account. That chapter will conclude with little support for an accusatory reluctance account, which claims the truth bias is an active self-derogation of accuracy. Instead, it seems social relatedness information can lead to judgmental biases in decision-making processes.

Summary: Socialisation practices. Over the years it has been suggested the truth bias may be a reflection of the social environment. To avoid conflict and social exclusion, accuracy may be sacrificed in favour of believing others. At the foundation of these accounts rests the idea that the perception of others as social beings leads to a reluctance to label people liars (DePaulo & Rosenthal, 1979; Ekman, 1992; Miller et al., 1986; O'Sullivan, 2003; O'Sullivan et al., 1988; Vrij, 2008; Vrij, Granhag &

Porter, 2010), known as an *accusatory reluctance*. Chapter 6 tests the account by examining whether the mere perception of others as social or non-social agents is sufficient to invoke or remove the truth bias, respectively.

An alternative account was also presented: social information is but another cue to aid decision-making. Rather than willingly sacrificing accuracy, raters may make use of the available social information as a simple heuristic to guide the judgment. These simple rules are often accurate and useful under uncertainty: trusting a friend over a stranger is in the long run a useful strategy. In the laboratory where the environment can be manipulated, these simple rules can be seen most readily. By giving the appearance that some speakers are psychologically closer to the rater than others, the biasing effects of this information can be observed. I begin to explore this account in Chapter 6. To pre-empt the results, the accusatory reluctance account is not supported. However, social relatedness can be used as a cue to guide the judgment: perceiving others as psychologically closer to us results in a tendency towards believing them. As in Chapters 4 and 5, the adaptive use of limited information in the environment can be used to guide the judgment.

Conclusion: Chapter 3

This chapter considered the type of information used by an adaptive decision maker. Key to the adaptive component of the account is that the selection of information is flexible and context-appropriate. I will argue that raters can make use of a variety of sources of information dependent on their availability and needs. In the absence of information, base-rate knowledge can be used. Having observed a set of behaviours, if uncertainty still persists then the uncertainty itself can be used to guide the judgment: truths are easier to spot than lies and so uncertainty can reflect deception. When available, knowledge about the social environment can also be brought to bear on the decision. This highlights the importance of understanding the social environment: it does not act independently, having only secondary effects on the decision process, but rather can be used to actively guide the decision.

Chapter 4: Testing the Dual Process Theory:

The Multi-Response Interview

Situations in which the bias is no longer present offer a window into studying the causal forces behind biased responding, and ultimately behind the decision-making process. In this chapter I will present four experiments. The first three experiments will test the challenges I issued against the HAM account in Chapter 2. They were (1) that the truth bias must be in part explained by the cognition of the rater independent of the behaviour of the speaker, (2) that the shift in bias over time must be the result of the amount of processing time available, and (3) that if it time cannot predict the change in bias, then the type of information available must. In the last experiment in this chapter will consider and show empirical support for an alternative account of the change in bias over time: raters are making use of a consistency heuristic.

Masip et al. (2006; Masip, Garrido, et al., 2009) noted that typical lie detection studies present brief snippets of behaviour, often no longer than 30 s in duration, which may be too short to give enough information to the rater and to give them sufficient processing time. In their study, they considered statements lasting over a period of minutes. To more closely simulate a police interview setup, each speaker was asked three questions about different aspects of the same mock crime scene they had just witnessed. They were instructed to either lie or tell the truth throughout their statement, meaning that each speaker produced either three deceptive or three truthful responses as part of their statement. On presenting recordings of these interviews to a set of rater participants, an initial truth bias was found after the first response, but by the second and third response the truth bias showed a significant reduction and their accuracy increased. The authors interpreted these findings as evidence of a default-interventionist HAM, which predicts an initial default heuristic processing is interrupted over time by a more analytical, less biased and less error-prone form of processing (Masip et al., 2006; Masip, Garrido, et al., 2009; Masip et al., 2010).

Whilst the findings are consistent with a HAM, there are a number of challenges that must be met if a HAM account is to be accepted over other potential explanations. First, the bias may be little more than an accurate reflection of the behaviour being rated (Fan et al., 1995; Zuckerman et al., 1987; Zuckerman, Koestner, et al., 1984). This challenge is taken up in Experiment 1. Evidence for an independent cognitive component to the truth bias will be shown.

Second, previous studies have only captured the judgment at the point of the first, second and third response of the statement. This offers a proxy for the amount of processing time, but because any given response can be long or short, it is not an accurate indication. Because both default-interventionist and parallel-competition models make a claim to the duration of processing time as the mediating factor between heuristic and analytical processing (Evans, 2007), it is necessary to more directly assess the effect of processing time on the truth bias. In Experiments 1 and 2, logistic mixed effects models will be constructed to test this prediction. Although the point of rating can predict the decline in truth bias, the duration of processing time as measured by the cumulative viewing duration up until the point of judgment cannot. This presents the first major challenge to the HAM.

If HAMs are to account for the truth bias, the switch between the two systems of thought cannot be determined by processing time because this was not supported in Experiment 1. Default-interventionist and parallel-competition class of models make this prediction (Evans, 2007), but a third and final class of models identified by Evans (2007), pre-emptive conflict resolution models, do not. Instead, they propose heuristic or analytical processing is chosen at the outset of the decision-making process. For the HAM to be supported, it must be shown that if processing time cannot account for the decline in the truth bias, the selection of heuristic or analytical processing at the outset of the judgment can instead. For example, the decline in the truth bias may be attributable to a switch from the use of visual cues to verbal cues. Where only visual or only verbal cues are present, there should be observed no decline in the bias across the ratings; rather, visual cues should show a consistent truth bias and verbal cues should show consistently unbiased responding. However, Experiment 2 fails to find support for this account, showing a decline in the truth bias regardless of the channel (visual or verbal) available.

In Experiment 3, the processing time challenge is revisited. Thus far, the mixed effects models have been used to statistically demonstrate the lack of an effect. A more robust test of the processing time prediction would be to manipulate the processing time. With shorter clips, there should be a lesser or no decline in the truth bias reflecting the continued use of heuristic processing, but with longer clips the decline in the truth bias should be observed as before. Consider also that the amount of processing time available ought be a useful cue as to whether to select heuristic or analytical processing from the outset, if a pre-emptive-conflict resolution model is true (Evans, 2007). However, in both conditions a declining truth bias is observed, again failing to support the HAM.

The decline in the truth bias can be observed across each new response, but it is not predicted by processing time. The presentation of each new response from the speaker may account for the decline in the truth bias. In Experiment 4, consideration is given to whether comparisons between the responses can account for the phenomenon. Note that this differs from the behavioural account tested in Experiment 1 insofar as it is the *perception* of consistency that is important, not whether the speaker truly is or is not consistent. Raters disagree as to whether a given statement is consistent or not (Granhag & Strömwall, 1999). The declining bias could reflect this perception. Experiment 4 supports this explanation, although it is noteworthy that consistency was found to be a diagnostic indicator of deception also.

Experiment 1: The Behavioural Account

Both truth-tellers and liars want to be believed to be telling the truth and so will adapt their behaviour accordingly (Buller & Burgoon, 1996; DePaulo et al., 2003; Gilovich et al., 1998). The decline in the truth bias and increase in accuracy over time may be an indication that liars become more nervous over time, for example, or leak cues to their deceit in other ways (see Granhag & Strömwall, 2002). To support a HAM account, it must be shown that there is a cognitive component to the truth bias that operates independently of the behaviour being displayed.

In line with the findings of Masip et al. (2006, 2009, 2010), the point at which the judgment is made (after the speaker's 1st, 2nd, or 3rd response) was expected to be a key determinant of the truth bias. In the studies reported throughout, a truth bias is measured by the proportion of truth judgments (PTJ) made, where a PTJ of greater than 0.5 indicates a truth bias.

Because each response can stand in isolation of the other responses in the statement, it is possible to reverse the order of the responses recorded across the interview. If liars became more nervous over time, for example, it may be expected that a reversal of the recording order would show an increase in the truth bias rather than a decrease. If, however, the declining truth bias reflects a cognitive shift on part of the rater, the order of the behaviours should make little difference to the pattern of responding over time. It was predicted that truth judgments would decline over successive ratings regardless of whether the speaker's first or last recorded response was presented initially. It was also hypothesised that accuracy would improve over successive ratings. These findings would offer initial support for a HAM account, where a shift towards an analytical process is indicated by less biased and less error-prone decision-making.

However, as discussed earlier, using the point of judgment may not be a useful proxy of time. Therefore, the *cumulative duration* of the speaker's responses until the moment of rating were also examined. It was predicted this duration would be negatively related to the PTJ, such that the PTJ would decline with longer viewing durations, and positively related to accuracy. As raters may shift towards an analytical mode of processing before the speaker finished the first response, the influence of the *duration of the speaker's first presented response* on raters' PTJ and accuracy was also assessed. Longer durations were expected to yield a smaller PTJ and greater accuracy rates.

To summarise the predictions, it was expected that for each new rating the truth bias would decline and accuracy would increase. To support a HAM position, the decrease in bias and increase in accuracy would have to be predicted by the amount of processing time available.

Method

Participants. Ninety-seven psychology undergraduates at the University of Salamanca volunteered to participate in exchange for an academic incentive. Fourteen participants (seven in each condition) indicated that they had taken a seminar on lie detection and were therefore excluded from analyses.[1] The final sample contained 83 participants (66 female; age M = 20.75, SD = 0.20, range: 18 to 35).

Materials. As described in the procedure section, I used a video set containing six speakers lying and six speakers telling the truth in response to three consecutive questions. These video clips had been recorded in an earlier study by Masip et al. (2006). A booklet was used to collect the respondents' binary truth-lie judgments, as well as their Likert-scaled confidence ratings, immediately after viewing each response.

Design. Raters were allocated randomly to one of two viewing conditions. Those allocated to the *direct viewing condition* (*n* = 44) watched the responses of each speaker in the order they were recorded, whereas those allocated to the *reverse viewing condition* (*n* = 39) watched the responses in reverse to their recording order. A stratified design was used in earlier work by Masip and colleagues (2009) to ensure that each experimental group contained approximately equal numbers of males and females with a similar distribution of ages. In this chapter the same stratified design

[1] Results reported throughout are similar with excluded participants added back into the analyses.

was used such that sex, χ^2 (1, N = 83) < 0.01, p = .995, and age, t (81) = 0.10, p = .920, d = 0.02, distributions did not differ substantially between the two experimental conditions.

The presentation order of each response (1st, 2nd, and 3rd presented response, within subjects), the viewing direction condition (direct or reverse, between subjects) and the veracity of each speaker (lies or truths, within subjects) served as independent variables. The dependent variables were the PTJ and the proportion of correct judgments (accuracy). Analyses of variance were run to examine the influence of presentation order vs. viewing direction on the PTJ and accuracy. Generalised logistic mixed-effects regression analyses were built in the R software package (R Development Core Team, 2011) using the LME4 (Bates, Maechler & Bolker, 2011) and languageR (Baayen, 2008) packages. The models were used to examine more closely the predictions made by default-interventionist and parallel-competition models. The specification of the models is detailed in the Results.

Procedure. *Mock crime stimuli.* The video stimulus set was adopted from Masip et al. (2006, 2009). To collect statements, Masip et al. showed speakers one of two videotapes. The narratives of each video were scripted such that minor details between videos changed, but the underlying structure was similar. Each video involved a woman, a man with a moustache and a man dressed in a suit, the latter of which acted as a clerk whilst the other two played customers. One of the customers asked to be served first because she or he was short of time. After the interview, this person inadvertently left a briefcase full of money on the desk. The second person, having noticed the briefcase, seized it and asked to be excused in order to check whether she or he had locked her or his car. The interviewer either allowed the thief to walk off, or noticed the briefcase being stolen and reprimanded the thief. The videos were later edited with the condition that they contained the same shots with the exception of the critical changes. The videos were over three minutes in length.

Speaker statements. After viewing the mock crime footage, speakers were interviewed. They were informed that they would be asked three questions about what the characters in the videotape did, that they would be asked the questions twice, and that in one interview they had to tell the truth, whilst in the other one they had to lie. They were given ten minutes to prepare their statements, although they did not know the specific questions to be asked. This is not untypical of police interviews, where defendants have time to consider their explanations but do not know what the interrogator may ask exactly. As motivation, speakers were told that observers would try to spot their lies. They were challenged to try hard to be convincing and were told that they would receive feedback—their names would be posted on a their classroom noticeboard scored for how easy or hard their lies were to spot.

The questions were always of the same format: "Describe in detail what the [man with a moustache/man in a suit/woman] did; I remind you that you have to [tell the truth/lie]". Thus, each speaker provided a single statement about a single event by responding to three questions, as in police interviews. The order of the questions, whether the speaker lied or told the truth first, and the selected mock-crime videotape were counterbalanced, with an additional aim of equally distributing males and females into each Question-order x Lie-order x Mock-crime-videotape cell. Later, the recordings of the 24 speakers were divided into four videos with 12 different speakers in each video. The same speaker never appeared lying and telling the truth in the same video. The speakers' responses were displayed in the clips, but in order for the video sets not to be too long the interviewer's responses had been removed.

Because of sample size limitations, just one video set was selected for the current study —the video set that achieved ratings most representative of the overall results in Masip et al.'s (2009) research (video set A1). Further details can be found in Masip et al. (2006). The first recorded response lasted on average 50 s, Response 2 averaged 37 s and Response 3 averaged 39 s. A one-way analysis of variance (ANOVA) indicated that these duration differences were not statistically significant, F (2, 35) = 0.26, p = .776, η^2 = 0.01.

Raters. Participants in the current study were randomly allocated either to the direct viewing condition (the responses of each speaker were shown in their original recording order; n = 44) or to the reverse viewing condition (the speaker's third recorded response was presented and judged initially, followed by the second recorded response, and then by the first one; n = 39). Differences in sample size were a result of several participants not showing up. The procedure replicated Masip et al.'s (2009), with the exception of this experimental manipulation. The data were collected in a classroom with a group of participants taking part at the same time (two sessions were held for each viewing direction condition). The participants were sat apart such that they could not see each other's responses. They were informed that each speaker provided only a single statement that consisted of three responses, and that a statement was either deceptive or truthful across the three responses. After each response, the video was stopped and the participants marked in the booklet their binary lie-truth judgment and their judgmental confidence (on a 1-to-7 scale, with higher ratings indicating more confidence).

Throughout this thesis raters were not given any information about the proportion of lies and truths they would encounter unless otherwise stated.

Results

Regardless of viewing order, the truth bias declined across successive ratings, supporting a role for the rater's cognition in the truth bias independent of the speakers' behaviours. However, the decline was not predicted by the amount of processing time available, evidencing the first challenge to the HAM position.

Throughout this thesis all unreported effects are non-significant.

Testing the behavioural account. Two 2 (Veracity: truthful/deceptive statement) x 3 (Presentation Order: 1st, 2nd, and 3rd presented response) x 2 (Viewing Condition: direct/reverse) ANOVAs with repeated measures on the former two variables were conducted, the first on the PTJ and the second on accuracy. The first ANOVA revealed a significant main effect of veracity on the PTJ, F (1, 81) = 16.55, $p < .001$, $\eta^2 = 0.17$, with truthful statements being more often judged truthful (M = .62, SD = .02) than deceptive statements (M = .53, SD = .02). The main effect of presentation order was also significant, F (1.62, 131.58) = 21.20, $p < .001$, $\eta^2 = 0.21$,[2] reflecting a decrease in truth judgments over time that was significant between the first (M = .61, SD = .16) and the second (M = .56, SD = .16) and between the first and the third (M = .54, SD = .17) response of the speakers (t (82) = 3.48, $p < .001$, d = 0.40, and t (82) = 4.08, $p < .001$, d = 0.46, respectively), but that was not significant between the second and third responses, t (82) = 0.60, p = .155, d = 0.08. This pattern replicates previous findings (Masip et al., 2006; Masip, Garrido, et al., 2009). The Presentation Order x Viewing Condition interaction was not significant, F (1.62, 131.58) = 2.71, p = .081, $\eta^2 = 0.03$, suggesting that the PTJ decreased over time

[2] Because viewing direction violated the assumption of sphericity, a Greenhouse-Geisser correction was applied.

regardless of the order in which the responses and their corresponding behaviours were presented. These findings support a cognitive influence on judgments. No other main effects or interactions were statistically significant in this ANOVA.

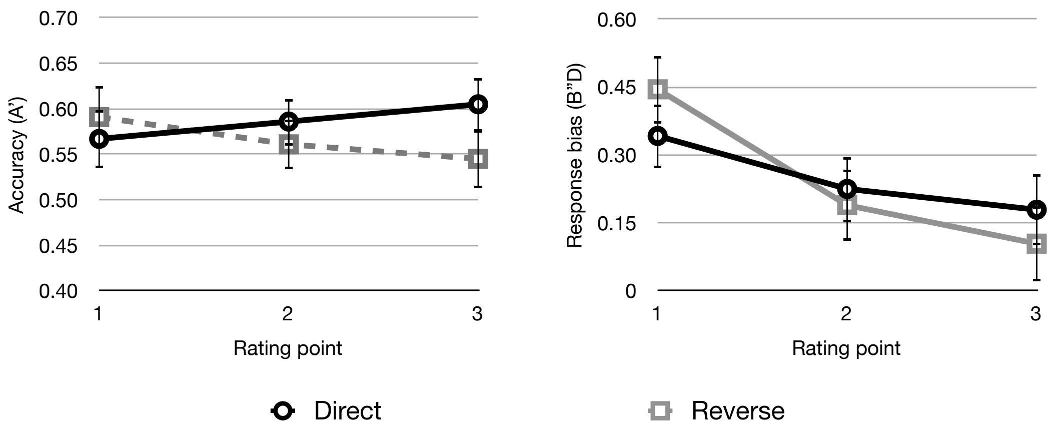

Figure 1. (a) Accuracy as measured by A’ across the three responses of the full statement split by viewing direction: either in the original recording order (direct) or in reversed order (reverse) (left). (b) Response bias as measured by B''_D across the three responses in the direct and reverse conditions (right). See in text for information on calculating and interpreting A’ and B''_D values. Whiskers denote standard error.

The second ANOVA revealed a main effect of veracity on accuracy, F (1, 81) = 20.64, $p < .001$, $\eta^2 = 0.20$. In line with a veracity effect (Levine et al., 1999), accuracy was greater for truths (M = .62, SD = .02) than it was for lies (M = .48, SD = .02). This main effect was moderated by the presentation order, F (1.62, 131.58) = 21.20, $p < .001$, $\eta^2 = 0.21$. The interaction revealed that, in judging truths, accuracy was higher for the first presented response than for the second, t (81) = 3.58, $p < .001$, d = 0.34, and third presented responses, t (81) = 4.05**,** $p < .001$, d = 0.37. In contrast, when judging lies accuracy was lower for the first presented response than the third presented response, *t* (81) = 3.79, $p < .001$, *d* = -0.34. No other main effects or interactions were significant in this ANOVA.

Note that a decrease in bias could cause an increase in accuracy: that is, there is shared variance between the variables. To separate out the independent effects of accuracy and response bias, signal detection measures of A', calculated using Rae's (1976) formula, and B''_D, calculated using Donaldson's (1992) formula, were used respectively. The accuracy measure A' is bounded between zero and one, with 0.5 equivalent to chance.

A 2 (Viewing Condition) x 3 (Presentation Order) ANOVA was conducted on the A' scores and found no significant main effects or interactions, all *p*s > .128 (see Figure 1a). B''_D measures response bias independent of accuracy effects and ranges from -1 (biased towards responding 'lie') to +1 (biased towards responding 'truth'), where zero is equivalent to no bias. The previous ANOVA was conducted again using B''_D as the dependent variable. A main effect of presentation order was found, F (1.53, 123.82) = 17.96, $p < .001$, $\eta^2 = .181$, such that the truth bias declined from the first (M = .39, SD = .47) to both the second (M = .21, SD = .48), t (81) = 4.16, $p < .001$, d = 0.38, and third rating (M = .14, SD = .52), t (81) = 4.85, $p < .001$, d = 0.50, but did not decline between the second and third ratings, t (81) = 2.10, p = .105, d = 0.14 (see Figure 1b). Thus the signal detection measures compliment the analyses conducted on the raw PTJ and accuracy scores.

Aside from the main effect of viewing order, which is an important prediction that has been met, it was also claimed that there was no interaction between the viewing condition (direct, reverse) and the presentation order, i.e. that the PTJ was not predicted by the first recorded, second recorded and third recorded behaviours. The difficulty with interpreting null *p* values and their associated statistics is that it is unclear whether the lack of finding an effect results from an underpowered study or type II error, or from the lack of a real effect. That is, a non-significant effect does not show support for the null hypothesis, but rather merely indicates a lack of evidence.

The Bayes factor circumvents this issue by asking how probable the data are given one model versus another. Data can be said to show support for or against the null or instead show the lack of an effect due to a lack of evidence in either direction. The BayesFactor package version 0.9.4 (Morey & Rouder, 2013) designed for the R statistical environment (R Development Core Team, 2011) was used to calculate a Bayes factor. A prior Cauchy scale of r = 0.5 over the effect sizes was selected; this prior includes 50% of the prior mass within the range of effect sizes between -0.5 and 0.5. This scaling factor is used throughout this thesis.[3]

A complex model entered the presentation order, the veracity of the speaker, the viewing condition, and their interactions as fixed effects, with the PTJ acting as the outcome variable. Fully specified random effects were included for rating participants and for speakers. This complex model was compared to a simpler model with the Presentation Order x Viewing Condition interaction removed. A Bayes factor indicated that in order to prefer the more complex model, we would need prior odds favouring it of greater than about 45. This was taken as strong support for the hypothesis that presentation order and viewing condition did not interact. The same analysis was conducted on the accuracy scores. The data were 100 times more likely

[3] A scaling factor of 0.5 is recommended by the BayesFactor documentation for most experimental designs given that it is readily computable and gives a stable integration of the likelihood. Lie detection research typically exhibits small effect sizes. It seemed appropriate to include 50% of the prior distribution's mass under an effect size of 0.5 for this reason. The average reported Cohen's *d* effect size in this thesis, for reference, is 0.48 (*SD* = .44).

under the null hypothesis, providing strong support for the lack of an interaction between presentation order and viewing condition.

In summary, the pattern of results was the same regardless of whether the statements were presented in the order in which they had been recorded or in the opposite order. The PTJ tended to decrease over successive judgments. Accuracy tended to increase for lies but decrease for truths indicating a shift towards responding 'lie' irrespective of whether raters were listening to a lie or truth.

Testing the default-interventionist and parallel-competition models. Having established that the pattern of rating is not driven solely by the behaviours being displayed, I asked whether heuristic processing strategies are at the root of the biased responding (and in turn the shift away from the truth bias). Because the time of judgment is not a robust proxy of viewing duration, the results so far provide only initial support for HAMs and are also consistent with the step-by-step response mode account. Default-interventionist models propose that cumulative viewing duration of a speaker should predict the PTJ (with a negative trend) and accuracy (with a positive trend), as the shift to analytical processing is dictated by time. In contrast, a step-by-step response mode account proposes cumulative viewing duration will not be able to predict the PTJ.

Because the data are not easily amenable to traditional F-tests, a model comparison approach was used to assess whether the addition of cumulative duration could significantly improve the model. Two mixed-effects generalised logistic regression models were created. The fixed effect of cumulative viewing time was added to a simpler model that included the veracity and the viewing condition to determine whether cumulative time could significantly improve the model fit. The simple model included the observer and the video-recorded speaker as random effects, each with its own random intercept. The random slopes included for the observer were the cumulative viewing direction, presentation order of the speakers' responses, and veracity of the statement. The random slopes for the speaker were the cumulative viewing time, viewing direction condition and the presentation order. That is, slopes for all variables were permitted provided a slope was possible to model (i.e. provided the speaker or observer could be found in more than one cell for the given variable), resulting in a maximally specified mixed effects model. Restricted maximum likelihood estimates of the models were based on the Laplace approximation. The PTJ and accuracy acted as outcome variables.

A significant difference in the predictive ability of these two models would indicate that the addition of cumulative viewing duration into the model significantly improved the fit of the data. It was found that the addition of cumulative duration did not significantly improve the model when attempting to predict either the PTJ, χ^2 (1) = 0.03, p = .861, or accuracy, χ^2 (1) = 1.22, p = .290. Thus the simpler model should be preferred.

It might be the case that analytical processing strategies were adopted early on, as some speakers provided particularly long statements. Indeed, one response was so long that it was greater than three standard deviations from the mean initial response duration and so could be considered outlier from the other initial responses.[4] Contrary to both default-interventionist and parallel-competition models, the duration of the first presented response could not significantly increase the prediction ability of either

[4] Similar results are found when this outlier is included back into the analysis, as is the case throughout this chapter.

the PTJ, χ^2 (1) < 0.77, p = .381, or accuracy, χ^2 (1) = 0.67, p = .411, in judging that response.

Discussion

In this study, raters judged the veracity of three successive statements of a number of speakers. Contrary to a behavioural account and consistent with the HAM, the PTJ was initially high and decreased over successive judgments regardless of the presentation order of responses. This finding suggests that the time of rating is a more potent influence on the veracity judgment than changes in behaviours that may occur over time.[5] The decrease in truth judgments resulted in an increase in accuracy in judging lies, but also in a decrease in accuracy in judging truths. As a result, overall accuracy did not increase over time; instead, raters were more likely to rate a statement as a lie irrespective of its actual veracity. This is contrary to predictions and is the first suggestive piece of evidence that raters were not switching to a more reasoned and analytical processing strategy, as is claimed by default-interventionist and parallel-competition accounts.

More conclusive evidence can be drawn from exploring the processing time raters had available. Contrary to these HAM models, regression analyses revealed that neither the cumulative viewing duration of a speaker's statement nor the duration of the initially viewed response could predict the PTJ or accuracy. This was surprising, as a proxy of viewing duration (the point of rating: after the 1st, 2nd, or 3rd viewed response) had a significant influence on both PTJ and accuracy. Thus, it appears that the act of rating over several occasions, irrespective of the length of a statement,

[5] Of course, I do not dismiss the important influence of the behaviours of the speaker: after all, a rater might be considered foolhardy to ignore an admission of deceit.

reduces the truth bias and may improve detection accuracy (overall accuracy did not increase in this study, but it did in Masip et al.'s, 2009).

However, these results do not allow us to dismiss HAMs altogether. A third and final class of HAM-based models identified by Evans (2007) outlined in the next section does not propose that analytical processing will only be seen later in the judgment process. This third class makes predictions about channel (visual, audio-visual, audio) effects on judgments and accuracy. Experiment 2 was conducted to test these predictions and determine whether HAMs can explain the presence of the truth bias and the decline in the PTJ across time.

Experiment 2: Channel Effects

According to HAMs, heuristic processing is a rapid process that requires little cognitive effort. Conversely, systematic processing is a slow, controlled and thoughtful process that involves considerable cognitive effort - and hence needs cognitive resources to take place (Chaiken, 1987; Chaiken et al., 1989; Chaiken & Stangor, 1987; Evans, 2008). This is so regardless of whether heuristic processing is assumed to run prior to systematic processing (Evans's, 2007, default-interventionist model), both heuristic and systematic processing are assumed to occur in parallel (Evans's, 2007, parallel-competitive model), or a selection is made at the outset between heuristic or systematic processing (this latter strategy is called *pre-emptive conflict resolution model* by Evans, 2007).

The selection can be motivated by internal factors such as task involvement (Chaiken, 1980; Chaiken et al., 1989; Chaiken & Maheswaran, 1994; Chaiken &

Trope, 1999; Chen & Chaiken, 1999; Forrest & Feldman, 2000; Petty & Cacioppo, 1979, 1986; Reinhard, 2010; Reinhard & Sporer, 2010), but can also be causally dependent on the type of information available in the external environment. Analytical processing is the preferred route when effortful processing is required (Chaiken, 1980; Evans, 2007; Evans et al., 1993). Speech content requires the listener to comprehend and reconstruct the narrative (Chen & Chaiken, 1999; Gilbert & Krull, 1988; Petty & Wegener, 1999). As such, processing speech requires greater cognitive effort and thus it is thought the analytical stream of processing is engaged (Forrest & Feldman, 2000; Gilbert & Krull, 1988; Reinhard & Sporer, 2008, Study 2; Stiff et al., 1989). Visual behaviours have been shown to require less cognitive effort to process (Forrest & Feldman, 2000; Reinhard & Sporer, 2008, Study 3; Stiff et al., 1989). When only visual information is present, the heuristic system is chosen (Reinhard & Sporer, 2008, Study 3; Stiff et al., 1989).

Thus, the channel of communication (i.e., the availability of visual vs. verbal information) should influence truth judgments and potentially accuracy. Specifically, visual cues should yield a disproportionate degree of truth responding, having been processed by heuristic modes of thought, whereas the availability of verbal information should result in a low or no truth bias. In addition, because analytical processing takes a systematic approach towards forming judgments accuracy should be higher when verbal cues (processed analytically) are available.

Gilbert and Krull (1988) showed precisely this; accuracy in interpersonal judgments was higher when using verbal cues if cognitive resources were available compared to when there was an additional cognitive load from a secondary task. Thus, provided no extraneous cognitive load is placed onto the raters, a heuristic account would predict that accuracy should be highest when verbal information is present and lowest when verbal information is lacking. Indeed, this is what extant meta-analyses show (Bond & DePaulo, 2006; DePaulo, Zuckerman & Rosenthal, 1980; Zuckerman et al., 1981; see also reviews by DePaulo, Stone & Lassiter, 1985; Vrij, 2008). In addition, verbal cues are more revealing of deception than nonverbal cues (DePaulo et al., 2003), and people have more accurate beliefs about verbal than nonverbal deception cues (Strömwall, Granhag & Hartwig, 2004; although see Hartwig & Bond, 2011, for evidence that raters do not necessarily use the cues they self-report as being diagnostic). Therefore, raters processing verbal cues should be more accurate than those with access to nonverbal information only.

The declining truth bias may be attributable to a switch in the use of cues, from visual cues to verbal cues. In Experiment 2, participants watched or heard the direct-viewing condition video of Experiment 1 and had to perform the same task. However, one third of participants had access to only visual information from the video (video condition), one third to both visual and audio information (audio-visual condition), and one third to audio information only (audio condition). In line with the dual-process account, it was predicted that more truth judgments would be made in the video than in the audio condition, with the audio-visual condition located between these, being the only condition to show an initially high truth responding that declines over time. Similarly, as predicted by the dual-process account, it was expected that accuracy would be lowest in the video condition and highest in the audio condition.

In addition, if the decrease in truth judgments is a result of a shift from heuristic to systematic processing as per default interventionist and parallel competition accounts, then the decrease would be weakest in the video condition because switching to systematic processing would be difficult with the unavailability of revealing verbal information. The decline in the PTJ should be strongest in the audio condition because the cues that are used for analytical processing (i.e. verbal cues) are available. Also, accuracy was expected to increase over consecutive judgments primarily in the audio condition, but not in the video condition for similar reasons consistent with a dual process position.

Method

Participants. Fifty-four psychology undergraduate students at a Spanish university volunteered to participate in exchange for an academic incentive and were allocated to either the video or the audio condition. Five participants (all in the video condition) indicated they had taken a seminar on lie detection and were therefore excluded from analysis. None of the remaining 49 participants (32 female; age M = 20.43, SD = 3.21, range: 18 to 38) had participated in Experiment 1. These participants took part in the audio only and video only conditions. The current experiment was set up to collect independent data from a third group who would have taken part in the audio-visual condition. However, technical problems arose when playing the clips that discontinued data collection from the first group. For this reason, the data from the first 24 participants in Experiment 1 who viewed the videos in the original recorded order (15 female; age M = 20.21, SD = 2.32, range: 18 to 25) were analysed here to allow a comparison with ratings made when full audio-visual information was available. These participants had attended the same session in Experiment 1. In total, 73 participants were analysed in this Experiment (47 female; age M = 20.36, SD = 2.94, range: 18 to 38).

Materials, design, and procedure. The booklet and stimulus video set used in Experiment 1 were employed. The design and procedure closely followed Experiment

1, except that here all participants viewed or heard the responses of each participant in the original recorded order and had access either to only the visual channel (*n* = 22), the audio channel (*n* = 27), or full audio-visual channels (*n* = 24). Sex, χ^2 (2, *N* = 73) = 0.20, *p* = .906, and age, *F* (2, 70) = 0.08, *p* = .928, distributions did not differ substantially between the three groups. The channel manipulation (visual, audio-visual, audio) served as an independent variable along with response number (1st, 2nd, or 3rd speaker's response) and the veracity (truthful, deceptive) of each statement.

Generalised logistic mixed-effects regression analyses similar to Experiment 1 were conducted to determine the effects of cumulative viewing duration. The dependent and outcome variables of these analyses were the PTJ (used to assess the extent of the truth bias) and accuracy.

Results

The hypotheses were generally unsupported: regardless of the information channel available, there was a decline in the truth bias over successive judgments. As in Experiment 1, no evidence was found that the amount of processing time could predict the decline in truth responding.

Truth bias. A 2 (Veracity: truthful/deceptive statement) x 3 (Response Number: 1st, 2nd, or 3rd response) x 3 (Channel: audio/visual/audio-visual) mixed ANOVA with repeated measures on the first two variables was run on the PTJ. The veracity main effect was not significant, *F* (1,70) = 2.27, *p* = .136, η^2 = .031. Replicating Experiment 1, a significant main effect of response number was present, *F* (1.74, 121.55) = 7.10, *p* = .002, η^2 = .092, showing a reduction in the PTJ over successive ratings (Figure 2). Bonferroni-corrected post-hoc *t*-tests found that the PTJ declined between the 1st (*M* = .58, *SD* = .14) and the 2nd response (*M* = .53, *SD* = .14), *t* (71) = 3.58, *p* = .002, *d* = 2.52, and between the 1st and the 3rd response (*M* = .53, *SD* = .16**),** *t* (71) = 2.80, *p* = .024, *d* = 0.33, but not between the 2nd and the 3rd response, *t* (71) = -0.09, *p* > .999, *d* < 0.01.

The critical channel main effect failed to reach significance, *F* (2, 70) = 2.70, *p* = .074, η^2 = .072. In fact, the means were in the opposite direction to what was predicted (for the video channel, *M* = .50, *SD* = .21; for the audio-visual channel, *M* = .57, *SD* = .22; for the audio channel, *M* = .58, *SD* = .23). Also, contrary to predictions the Response Number x Channel interaction was not significant, *F* (3.47, 121.55) = 1.10, *p* = .357, η^2 = .030. The reduction in the PTJ was very meagre for all three channels. However, as predicted, it was weakest in the video condition, for which none of the three planned pairwise comparisons were significant, all *p*s > .519.

A significant decrease was apparent for the audio channel between the 1st (*M* = .61, *SD* = .23) and the 2nd response (*M* = .55, *SD* = .22), *t* (25) = 3.25, *p* = .005, *d* = 0.27, but not between the 2nd and the 3rd response (*M* = .56, *SD* = .26), *t* (25) = -0.63, *p* > .999, *d* = -0.04; for the comparison between the 1st and the 3rd response, *t* (25) = 2.08, *p* = .124, *d* = 0.20. The PTJ decreased with marginal significance for the audio-visual channel (means and standard deviations for the 1st and 3rd response were *M* = .60, *SD* = .24, and *M* = .54, *SD* = .27, respectively), *t* (20) = 2.33, *p* = .067, *d* = 0.24, although the decrease from the first to the second response (*M* = .57, *SD* = .23) or from the second to the third one were not significant, *t* (20) = -1.66, *p* = .312, *d* = 0.13, and *t* (20) = 1.40, *p* = .494, *d* = 0.12, respectively.

Signal detection measures confirmed a main effect of response number on bias, B''_D, *F* (1.75, 122.22) = 6.90, *p* = .002, η^2 = .090, with no other main effects or interactions. Bonferroni-corrected pairwise comparisons indicated a decline between the first (*M* = 0.28, *SD* = 0.47) and both the second (*M* = 0.13, *SD* = 0.49), *t* (71) = -3.18, *p* = .007, *d* = 0.31, and third (*M* = 0.13, *SD* = 0.52) responses, *t* (71) = -2.85, *p* = .018, *d* = 0.30, but not between the second and third response, *t* (71) = -0.14, *p* > .999, *d* < 0.01 (see Figure 2b).

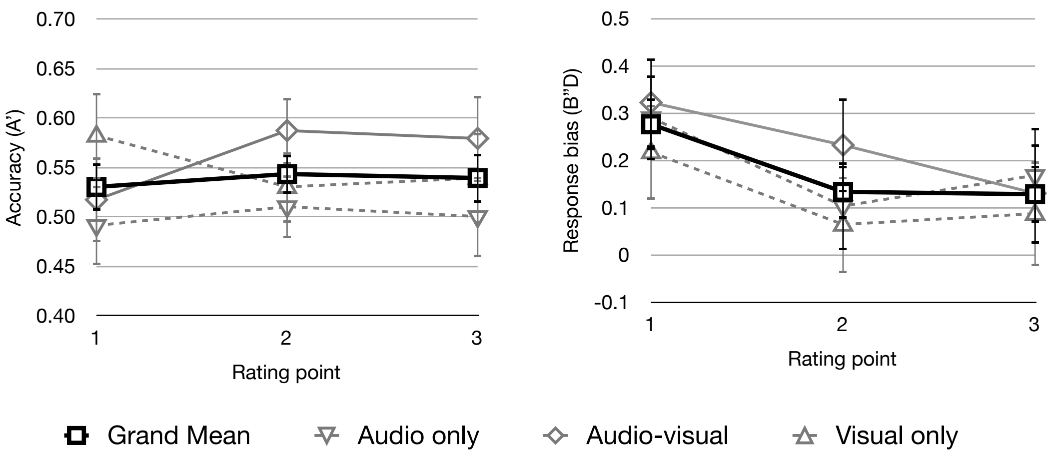

Figure 2. (a) Accuracy as measured by A' across the three responses of the full statement, whether given access to video only (dashed grey line, triangle), audio only (dashed grey line, inverted triangle) or the full audio-visual behaviours (solid grey line, diamond). Estimated marginal means for the overall pattern of responding is shown in black (left). (b) Response bias as measured by B''_D across the three responses in the audio only, visual only or audio-visual conditions, with the overall pattern shown in black (right). Whiskers denote standard error.

Bayes factors were calculated to test a main effect of channel and a Channel x Response Number interaction using a Cauchy prior with a scaling factor of r = 0.5 (Morey & Rouder, 2011). A full model with channel, speaker veracity and response number (and their interactions) were compared with a simpler model that did not include the channel main effect, both of which included fully specified random effects for rating participant and speakers. The data were 3.3 times more likely under the null hypothesis, supporting the lack of a main effect of channel on the PTJ. A second Bayes factor compared the full model to a simple model that removed the Channel x Response Number interaction. The data were about 950 times more likely under the simpler model, providing strong evidence against the role for an interaction effect.

Again replicating Experiment 1, a generalised logistic mixed effects regression analysis with maximal random effects determined that cumulative viewing duration could not add any predictive value to the model in fitting the PTJ, χ^2 (1) = 0.07, p = .790, in either the audio only, χ^2 (1) = 0.08, p = .772, or in the visual only Channels, χ^2 (1) = 0.17, p = .680. However, raters may have switched from a heuristic to a more systematic processing method even before the end of the speaker's first response. Overall there appears to be no effect of cumulative duration on the PTJ to the first response, χ^2 (1) = 0.82, p = .364, replicating Experiment 1. This was true whether only audio, χ^2 (1) = 0.82, p = .366, or only video information was present, χ^2 (1) = 1.11, p = .293, standing in contrast to both a default-interventionist and a parallel-competition account of heuristic processing.

Accuracy. A 2 (Veracity: truthful/deceptive statement) x 3 (Response Number: 1st, 2nd, or 3rd response) x 3 (Channel: audio/visual/audio-visual) mixed ANOVA with repeated measures on the former two variables was run on accuracy. The only significant main effect was for veracity, F (1, 70) = 14.90, p < .001, η^2 = .175, such that raters were more accurate in judging truths (M = .58, SD = .02) than lies (M = .46, *SD* = .02).

This was moderated by Response Number, F (1.73, 121.43) = 8.81, p = .001, η^2 = .112. Accuracy for lies increased between the 1st (M = .43, SD = .21) and the 2nd response (M = .48, SD = .17), t (71) = -3.16, p = .007, d = -0.26, and between the 1st and the 3rd response (M = .48, SD = .20), t (71) = -2.52, p = .047, d = -0.24, although it did not shift between the 2nd and 3rd response, t (71) = 0.25, p > .999, d < 0.01. Accuracy for truths decreased over successive judgments, but not significantly (for the 1st response, M = .61, SD = .02; for the 2nd response, M = .57, SD = .02; for the 3rd response, M = .56, SD = .21). That is, the judgments of truth decreased, slightly decreasing accuracy in judging truths and increasing accuracy in judging lies. Neither the channel main effect, F (2, 70) = 0.85, p = .431, η^2 = .024, or the Channel x Response Number interaction, F (3.57, 125.06) = 0.782, p = .526, η^2 = .022, were significant, failing to support the prediction that accuracy would be dependent on the type of information channel available.

An ANOVA conducted on A' scores, a measure of accuracy independent of bias, found no significant main effects of response number or channel, nor any significant interaction between them, all *p*s > .162 (see Figure 2a).

Bayes factors were calculated to test the main effect of channel on accuracy scores and the interaction of channel with response number on accuracy. A full model with all main effects and interactions between channel, response number and speaker veracity predicting accuracy was compared against a simpler model removing the channel main effect, both of which included fully-specified random effects for rating participant and speakers. The data shift the relative plausibility of the alternative to the null model by a factor of around 14, supporting the null. The same analysis was recomputed, except the simple model removed only the Channel x Response Number interaction. In order to prefer the alternative model, we would need prior odds favouring it of greater than about 1000, again favouring the null.

A comparison of generalised linear mixed effects models that either made use of or did not make use of cumulative duration as a predictor found that, contrary to a default-interventionist model, cumulative viewing duration could not predict accuracy, χ^2 (1) < 0.02, p = .878, in either the audio, χ^2 (1) < 0.01, p > .999, or video conditions, χ^2 (1) = 0.04, p = .840. Again, the duration of the first portion of the statement could not predict accuracy when rating the first response, χ^2 (1) = 2.12, p = .145. If only audio information was available, the duration of the first response could predict accuracy to it, χ^2 (1) = 5.90, B = 0.01, p = .015, but not if only video information was presented, χ^2 (1) = 1.21, p = .272. Provided only audio information was available, with greater cumulative duration of the initial response came improved accuracy. This offered partial support for the predicted increase in accuracy with longer viewing times.

Discussion

Overall, the effects found in Experiment 1 were replicated here; there was little support for the notion that the initial truth bias demonstrated in Experiment 1 and 2 resulted from heuristic processing. Accuracy and truth judgments were similar across all three conditions, and changes over time did not match HAM-based predictions. In addition, the total amount of time available to the rater to process the information was not predictive of their bias at either the end of the speaker's first response or over all three responses regardless of modality. Equally, accuracy could not be predicted from cumulative viewing time, with one exception: provided only audio information was available, the duration of the first response was able to predict accuracy such that accuracy increased with longer processing times.

This latter finding is consistent with an HAM: accuracy should increase with viewing time provided there are diagnostic cues available to the rater. Research shows raters are typically more accurate when making judgments from verbal than from nonverbal cues and that the former offer more valid indicators of deceit than the latter (see Sporer & Schwandt, 2007; Vrij, Granhag & Porter, 2010). However, HAMs explicitly predict a reduction in bias when switching from a heuristic to more analytical processing (Evans, 2007), which was not supported across the two studies reported thus far.

One possible explanation for this effect rests on the serial nature of speech. Whilst there is relatively greater diagnosticity in speech content than visual behaviour, time is required for the speech to unfold and the diagnostic cues to become available. Indeed, cues such as latency and amount of time spent talking that offer some degree of diagnosticity (DePaulo et al., 2003; Sporer & Schwandt, 2006) are by definition time dependent. Thus as time progresses, there is the potential for greater numbers of cues to be enumerated (Masip et al., 2006; Masip, Garrido, et al., 2009). In contrast, visual behaviours can be presented in parallel (e.g. averted gaze and reduced overall bodily movement); even if visual cues were more diagnostic than speech content, greater viewing times may not lead to a greater enumeration of valid indicators.

This account of increased accuracy in the audio only condition carries three caveats. First, this explanation is clearly post-hoc and speculative; future research could seek to exploit these differences between the channels of communication and explore how serial (speech cues) and parallel (visual behaviours) information is processed by the lie detector.

Second, it should be noted in the audio only condition raters had access to paraverbal as well as verbal cues. Paraverbal indicators such as pitch have been shown to accurately distinguish liars from truth-tellers (DePaulo et al., 2003), and thus speech may also be considered to offer parallel information. Assuming raters used both the verbal and paraverbal cues available to them in the audio only condition, this would suggest the distinction between serial and parallel processing is not as important as the distinction that speech cues offer more diagnostic information than visual cues.

Finally, whilst viewing time of the first response could predict accuracy in the audio only condition, this was not true of the audio-visual condition. This suggests the presence of visual information can actively impede the potential for improved lie detection accuracy, as found in prior research (Bond, Howard, Hutchison & Masip, 2013; Maier & Thurber, 1968; Stiff et al., 1989). A recent study found when a perfectly diagnostic cue to a speaker's intention to deceive was available, it was utilised with near perfect accuracy. However, when the raters were also shown the visual (but not the verbal or paraverbal) behaviour of the speaker, accuracy dropped markedly from 97% to 76% (Bond et al., 2013, Study 3).

Thus far processing time has been unable to predict the decline in response bias. In order to test the effects of processing time more directly, in Experiment 3 the duration of processing time available to the rater was explicitly manipulated. The findings presented thus far suggest the manipulation should have no significant effect on the response bias, whilst default-interventionist and parallel-competitive HAMs would predict a decline in the PTJ with an increase in the available processing time.

Experiment 3: Thin-Slicing

Whether accurate impressions about others can be formed from snippets of behaviour is unclear (Carney, Colvin & Hall, 2007). Some findings indicate accurate judgments of others can be greater with only short 'thin slices' of behaviour of between 5 and 20 s compared with judgments made using the full behavioural repertoire (Albrechtsen, Meissner & Susa, 2009; Fowler, Lilienfeld & Patrick, 2009). Others find similar accuracy rates whether viewing short or longer clips (Ambady & Rosenthal, 1992, 1993). Others yet, perhaps unsurprisingly, find that with greater amounts of information comes improved accuracy within limits, where accuracy gains eventually tail off (Borkenau, Mauer, Rieman, Spinath & Angleitner, 2004; Carney et al., 2007). I do not know of any studies that explicitly consider the effect of thin slicing on response bias, although Albrechtsen et al. (2009) report data showing no effect of thin slicing on the degree of biased responding.

Nonetheless, in accord with the predictions of default-interventionist and parallel-competition HAMs short viewing durations of 8 s of behaviour was predicted to result in a greater reliance on heuristic processing and thus be more heavily truth biased compared to full-length clips of behaviour.

Method

Participants. One hundred twenty-one undergraduate students took part in exchange for course credit. Thirty-nine participants indicated they had taken classes on lie detection and were excluded. This left 82 participants (63 female, age M = 19.29, SD = 3.22, range 18 to 36 years).

Materials, design, and procedure. The booklet and stimulus video set used in Experiments 1 and 2 were employed. The design and procedure closely followed Experiment 1, except here all participants viewed or heard the responses of each participant in the original recorded order and either viewed the full-length responses of each speaker (*long clips* condition, $n = 49$) or saw only the first 8 s of each speaker's response (*short clips* condition, $n = 33$). That is, those in the short clips condition saw all three responses made by each speaker, but only saw the initial 8 s of those responses. As mentioned in the methods section of Experiment 1, the first recorded response of the long clips lasted on average 50 s, Response 2 averaged 37 s and Response 3 averaged 39 s. That is, those in the long clips condition saw the full-length clips shown to participants in Experiment 1. Note that the statements were of varying length for the long clips condition but in the short clips condition all statements were held at a constant 8 s duration. Sex, χ^2 (1, $N = 82$) < 0.52, $p = .470$, and age, F (1, 80) = 0.09, $p = .926$**,** distributions did not differ substantially between the two conditions. The independent variables were clip length (long, short), response number (1st, 2nd, or 3rd speaker's response) and the veracity (truthful, deceptive) of each statement. The PTJ and accuracy were the dependent variables.

Results

Whether raters had a relatively long or short period of time in which to process the information had no effect on the declining truth bias, offering direct experimental support for the notion that processing time cannot predict the decline in the PTJ.

Truth bias. A 2 (Veracity: truthful/deceptive statement) x 3 (Response Number: 1st, 2nd, or 3rd response) x 2 (Clip Length: short or long) mixed ANOVA with repeated measures on the first and third variables was run on the PTJ. There was a main effect of veracity, F (1, 80) = 18.32, $p < .001$, $\eta^2 = 0.19$. Truthful statements were more often judged to be true (M = .59, SD = .16) than were deceptive statements (M = .49, SD = .18). There was also a main effect of response number, F (1.73, 138.41) = 9.12, $p < .001$, $\eta^2 = 0.10$, replicating Experiments 1 and 2. Bonferroni-corrected post-hoc *t*-tests indicated a significant decline from the first (M = .57, SD = .15) to the second response (M = .53, SD = .14), t (81) = 2.86, p = .013, d = 0.28, a difference that was still significant at the point of the third response (M = .51, SD = .15), t (81) = 3.50, p = .002, d = 0.40. There was no significant decline between the second and third response, as in Experiments 1 and 2, t (81) = 1.45, p = .411, d = 0.14.

The response number interacted with veracity, F (1.85, 148.26) = 9.49, $p < .001$, $\eta^2 = 0.10$, such that the PTJ for deceptive statements showed a decline from the first (M = .55, SD = .22) to the second response (M = .48, SD = .21), t (81) = 3.24, p = .005, d = 0.33, and from the second to the third response (M = .43, SD = .19), t (81) = 3.13, p = .009, d = 0.25, whilst for truths no such decline in the PTJ was evident at any point of rating, all *p*s > .890 (M = .59, .58 and .60, SD = .19, .17 and .21 for the first, second and third response, respectively). This interaction is indicative of an increase in accuracy: the proportion of truth responses declined in response to lies, but not in response to truths, over successive ratings of an individual.

While there was no main effect of clip length, F (1, 80) = 1.99, p = .163, η^2 = 0.02, there was a statistically significant Clip Length x Response Number interaction, F (1.73, 138.41) = 3.55, p = .038, $\eta^2 = 0.04$. There was no significant decline in the PTJ for the short 8 s clips at any point of rating, all *p*s > .999, see Figure 3. However, there was a significant decline in the PTJ when viewing the full length clips from the first to the second response, t (81) = 3.71, p = .001, d = 0.44, with the difference between the first and third response also significant, t (81) = 4.60, p < .001, d = 0.62, but as in Experiment 1 and 2 no significant decline from the second to the third response, t (81) = 2.07, p = .123, d = 0.21.

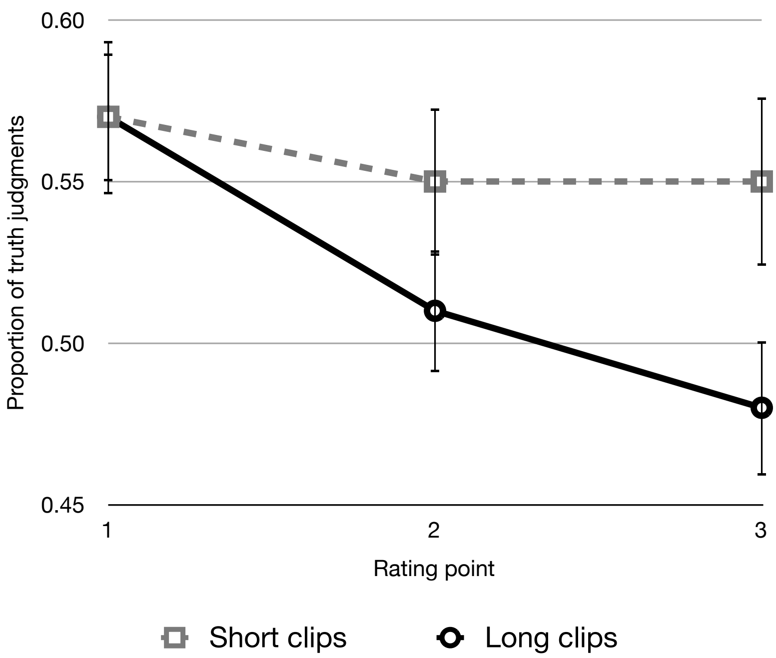

Figure 3. The shift in the PTJ over successive ratings split by whether raters were shown 8 s thin slices of each response or the full length response. Whiskers denote standard error.

Finally, there was a significant three-way interaction between veracity, response number and clip length, F (1.85, 148.26) = 6.93, p = .002, η^2 = 0.08. For truthful long clips, the typical pattern of declining truth judgments was observed. Truth judgments declined from the first (M = .62, SD = .19) to the second response (M = .54, SD = .17), t (48) = 3.42, p = .003, d = 0.44, but did not decline further at the point of the third response (M = .55, SD = .20), t (48) = -0.33, p > .999, d = -0.05. Deceptive long clips also showed an overall decline in the PTJ. The PTJ at the point of the first response (*M* = .51, *SD* = .21) was not significantly greater than at the second response (*M* = .47, *SD* = .20), *t* (48) = 1.63, *p* = .300, *d* = 0.20, but did decline from the second to the third response (*M* = .41, *SD* = .19), *t* (48) = 3.10, *p* = .008, *d* = 0.31. Whether listening to lies or truths, there appears to be a general decline in the PTJ across the response numbers.

For truthful statements in the short clips, there was an increase in the PTJ from the first (*M* = .56, *SD* = .18) to the second response (*M* = .62, *SD* = .17). This increase was not significant, *t* (32) = -2.00, *p* = .146, *d* = -0.34, but exhibited a medium effect size (Cohen, 1988). This increase continued to the third response (*M* = .65, *SD* = .20), and whilst the difference between the second and third response number again was not significant, *t* (32) = -1.08, *p* = .841, *d* = -0.16, there was an overall significant increase between ratings made at the point of the first and third response, *t* (32) = -2.58, *p* = .037, *d* = -0.47. Deceptive statements, in contrast, showed a significant decline from the first (*M* = .58, *SD* = .22) to the second response (*M* = .49, *SD* = .21), *t* (32) = 2.84, *p* = .019, *d* = 0.42, but no significant decline between the second and third response (*M* = .45, *SD* = .19), *t* (32) = 1.40, *p* = .502, *d* = 0.20. Thus when rating short clips, the trend appears to be a decline in the PTJ for deceptive statements and an increase in the PTJ for truthful statements, which amounts to increasingly accurate judgments across the speakers' responses.

That is, the three-way interaction showed an overall decrease in the truth bias for long clips, but for short clips showed an overall increase in accuracy.

Accuracy. A second 2 x 3 x 2 ANOVA was conducted as above using accuracy scores as the dependent variable. There was a main effect of response number, *F* (2,

160) = 5.02, p = .008, η^2 = 0.06. Bonferroni-corrected pairwise comparisons indicated a significant increase from the first (M = .52, SD = .14) to the second response (M = .56, SD = .14), t (81) = -3.23, p = .004, d = -0.29, but accuracy at the third response (M = .54, SD = .12) did not significant differ from accuracy at either the first or the second response, ps > .214. This main effect was moderated by clip length, F (2, 160) = 4.37, p = .014, η^2 = 0.05, see Figure 4. For long clips, accuracy was greater when rating the third response compared to the first response, t (48) = -2.82, p = .022, d = 0.29, but accuracy at neither of these points significantly differed accuracy in rating speakers' second responses, all ps > .362. For short clips, accuracy was greater when rating the second response compared to both the first response, t (32) = -2.95, p = .011, d = -0.44, and the third response, t (32) = 2.71, p = .027, d = 0.50, with accuracy at the first and third responses not significantly different, t (32) = -0.10, p > .999, d < 0.01. There was also a marginally significant effect of veracity, F (1, 80) = 3.56, p = .063, η^2 = 0.04. Raters were more accurate in judging truthful (M = .56, SD = .13) than deceptive (M = .52, SD = .18) statements, exhibiting a veracity effect (Levine et al., 1999).

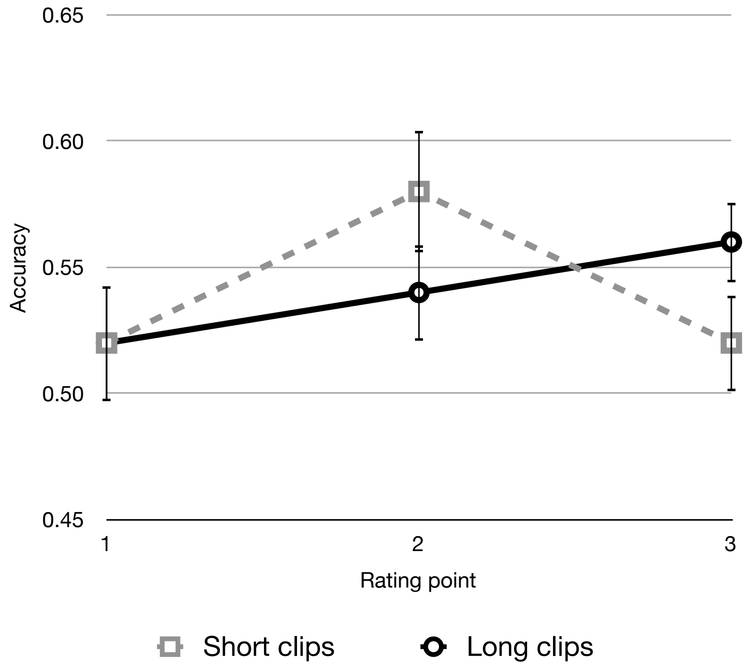

Figure 4. Proportion correct across successive ratings for those who judged 8 s thin slices and for those viewing full-length clips. Whiskers denote standard error.

Finally, there was a significant Veracity x Response Number interaction, F (2, 160) = 22.77, $p < .001$, $\eta^2 = 0.22$. Accuracy for truthful statements did not differ significantly between the first (M = .58, SD = .17) and second response (M = .60, SD = .21), t (81) = -1.06, p = .891, d = -0.11, although accuracy at both the first and second responses was significantly greater than at the point of the third response (M = .51, SD = .07), t (81) = 3.78, p = .001, d = 0.54, and t (81) = 3.86, p = .001, d = 0.58, respectively. Accuracy for lies increased from the first (M = .46, SD = .22) to the second rating (M = .52, SD = .21), t (81) = -3.24, p = .005, d = -0.32, and increased further by the point of the third response (M = .57, SD = .19), t (81) = -3.13, p = .009, d = -0.25.

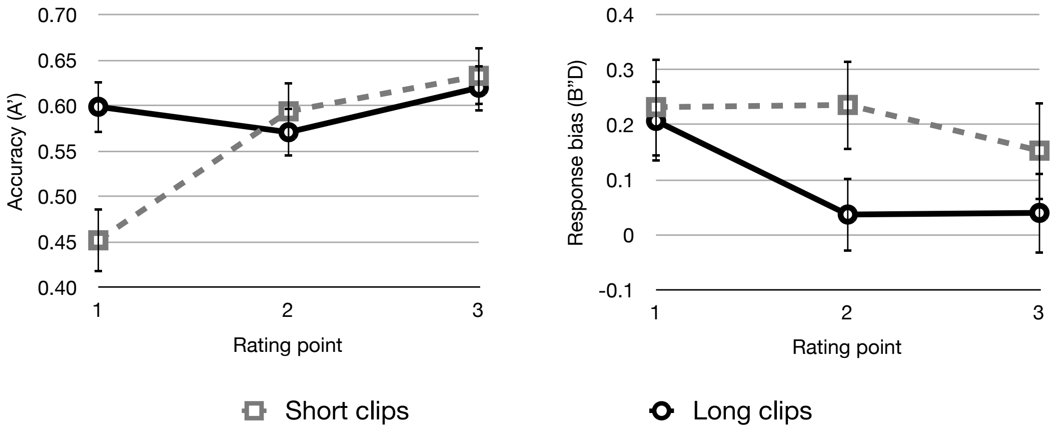

Figure 5. (a) Accuracy as measured by A' across the three long or short responses of the statement (left). (b) Response bias as measured by B''_D across the three long or short responses of the statement (right). Whiskers denote standard error.

Signal detection measures. The findings presented so far illustrates the importance of exploring the independent effects of response bias and accuracy on judgments, where interactions present in the first ANOVA exploring the PTJ made a complimentary appearance in the second ANOVA exploring accuracy scores. A 3 (response number) x 2 (clip length) ANOVA was conducted on the B''_D scores to examine response bias independent of accuracy. A main effect of response number was found, F (1.62, 129.58) = 4.63, p = .017, η^2 = 0.06. Response bias showed an overall decline form the first (M = .22, SD = .52) to the third response (M = .06, SD = .52), t (81) = 2.51, p = .042, d = 0.81, although bias at the point of the second response (M = .14, SD = .48) did not differ significantly from response bias at either the first, t (81) = 1.63, p = .322, d = 0.15, or the third response, t (81) = 1.86, p = .195, d = 0.15. There was no significant main effect of clip length, F (1, 80) = 2.18, p =

.144, $\eta^2 = 0.03$, or Response Number x Clip Length interaction, $F(1.62, 129.58) = 1.69$, $p = .193$, $\eta^2 = 0.02$. Thus regardless of clip length, whether long or short, response bias was seen to decline across the course of a speaker's statement (Figure 5b), as indicated by the main effect.

A Bayes factor was calculated using a Cauchy prior distribution with a scaling factor of r = 0.5. A simple model included all the variables associated with the preceding ANOVA with the interaction term between the response number and clip length removed, both models entering fully specified random effects. In order to prefer the more complex model, we would need prior odds favouring it of greater than 2.9, offering moderate support for the null hypothesis of no interaction effect.

A similar 2 x 3 ANOVA was conducted on the A' scores to explore accuracy. Again, a main effect of response number was found, $F(1.97, 157.45) = 10.96$, $p < .001$, $\eta^2 = 0.12$, which interacted with clip length, $F(1.97, 157.45) = 9.81$, $p < .001$, $\eta^2 = 0.11$. Accuracy did not change significantly between the first ($M = .60$, $SD = .20$), second ($M = .57$, $SD = .19$) and third responses ($M = .62$, $SD = .18$) when rating long clips, all *p*s > .181. For short clips, there was an increase in accuracy from near chance rates at the point of the first response ($M = .45$, $SD = .20$) to the second response ($M = .59$, $SD = .19$), $t(32) = -4.06$, $p < .001$, $d = -0.72$, but did not increase further between the second and third response ($M = .63$, $SD = .18$), $t(32) = -1.22$, $p = .665$, $d = -0.22$. Thus the interaction indicates relatively low accuracy after watching an 8 s clip with higher and relatively stable accuracy rates with longer viewing periods (whether in the long or short clip condition: Figure 5a).

In summary, a significant decrease in response bias was found across rating points, but the length of the clip did not moderate this. Neither was there a main effect of clip length on response bias. However, clip length did appear to influence accuracy rates, such that with only 8 s of viewing mean accuracy was below .50 and with longer viewing durations accuracy increased to approximately .60.

Discussion

Default-interventionist and parallel-competition models propose a greater reliance on heuristic processing when only given a short period of time with which to process information. As such, if they were to account for the pattern of declining truth bias across a speaker's statement, a greater degree of truth bias would be predicted when viewing short clips compared to when viewing full-length recordings of the statement. However, consistent with the mixed effects models conducted in Experiments 1 and 2, no support is found for the contention that truth bias declines with viewing time. Regardless of whether viewing long or short clips, there was a decline in truth bias, replicating Experiments 1 and 2 as well as prior research (Masip et al., 2006; Masip, Garrido, et al., 2009).

Unlike the first two experiments presented here, there was an increase in accuracy between the first and second response in the short clip condition. Accuracy improved from around chance rates after 8 s of viewing time to 60% with longer viewing durations. Thus one plausible account is that there simply was not enough delineating information in 8 s of viewing that allowed for any accurate classification of lies and truths.

So far the behavioural account of the truth bias has been tested (and rejected): regardless of the order in which behaviours were presented I found an initially high truth bias that declined across successive responses provided by a given speaker. I then turned to an explanation first suggested by Masip et al. (2006): HAMs can account for the decline with greater processing time. In three experiments support was not found for three general classes of HAMs identified by Evans (2007). Viewing time does not appear to predict the declining truth bias, but rather the number of responses made. Consideration is given to how the presentation of multiple responses across the statement gives rise to the possibility for comparison.

Experiment 4: Consistency Heuristic

The act of making a rating is itself an influence on what judgment is reached (Granhag & Strömwall, 2000b; Hogarth & Einhorn, 1992). In thinking about how new judgments may be influenced by prior judgments, consideration was given to how it might be possible to account for the decline in truth bias between the first and second response, but with no further decline between the second and third response. This finding has been evidenced in two previous studies (Masip et al., 2006; Masip, Garrido, et al., 2009) as well as consistently exhibited in Experiments 1 through 3, regardless of any experimental manipulation.

The act of rating after each new response provides not only new information but also new affordances for judgment. When two or more responses have been provided, it is possible to make comparisons between them. When the opportunity for comparison arises, raters utilise consistency of the tale more so than any other cue, such as the amount of detail they provided and even the plausibility of the statement (Granhag & Strömwall, 1999, 2000a, 2001b; see also Strömwall et al., 2003). Consistency appears to have low diagnostic utility in distinguishing adults' lies from truths (Granhag & Strömwall, 1999, 2001a, 2002; Strömwall et al., 2003; although see Strömwall & Granhag, 2007, for evidence of consistency as a predictive cue of children's deceptions).

Consistency seemed a plausible candidate for explaining the decline in the initial truth bias over time. Because raters perceive inconsistencies even when they are not present, there could be a shift towards perceiving others as deceptive when the opportunity for comparison arises, i.e. after having viewed the second response. Having established inconsistency after the second response, there may be no additional effect of continued perceived inconsistency by the third response. Thus it was predicted speakers who were considered more inconsistent would be rated such that a greater decline in the PTJ between response one and two would be observed compared to speakers who were considered more consistent.

Method

Participants. Forty-nine undergraduates (40 females, age M = 19.04, SD = 2.78, range 17 to 30 years) rated the stimulus videos for verbal and nonverbal consistency in part fulfilment of their course. To determine whether consistency was able to account for the decline in the PTJ, the lie-truth judgments collected from those studies in which the full behavioural repertoire was available to raters (i.e. Experiment 1, n = 83, and the Long Clips condition of Experiment 3, n = 49) were collated along with those published in Masip et al. (2009) that rated the same video stimulus set (n = 14). This yielded data from 146 participants (age M = 20.35, SD = 2.76, range 18 to 36 years) who viewed the full behavioural repertoire, 117 (80.1%) of which were female.

Materials and procedure. An instruction sheet provided definitions of verbal and nonverbal consistency. Verbal consistency was defined as 'how often the same or similar details are repeated over the three parts of their statement without contradicting themselves'. Nonverbal consistency was defined as 'how often the same or similar behaviour is repeated over the three parts of their statement'. The booklet requested two consistency ratings – one for verbal and another nonverbal consistency – of speakers between the first and second response. Primary interest was in accounting for the decline in the PTJ between the first and second response in earlier studies. However, all three responses from a speaker were presented to as closely as possible replicate the setup of prior studies. So that this third statement served a purpose and did not appear out of place, participants were also required to indicate how consistent speakers appeared across all three responses. Thus for each speaker, there were four responses to be made: verbal and nonverbal consistency across the first and second response, and across the first, second and third response.

The video stimuli from Experiment 1 direct condition was used for rating.

Design and data collation. Verbal and nonverbal consistency ratings were collected for each speaker after viewing two of the speakers' responses and after viewing all three responses, giving a 2 (veracity: lie or truth) x 2 (point of rating: after the second or third response) within-subjects design. These ratings were used initially to determine whether consistency was a diagnostic indicator of deception.

Each cell of the 2 (channel: verbal, nonverbal) x 2 (point of rating: after the second or third response) consistency ratings array collected about the 12 speakers was used to median split the speakers. This resulted in a high/low split for each cell of the 2 x 2 array. Incorporating the veracity of the speakers, this gave a 2 (verbal consistency: high, low) x 2 (nonverbal consistency: high, low) x 2 (point of rating: second or third) x 2 (veracity: lie, truth) within-subjects design.

Cronbach's alpha values were calculated to determine whether collapsing across channel (verbal, nonverbal) would be appropriate. Thus for each of the 2 (consistency: high, low) x 2 (consistency until rating point: second, third) x 2 (veracity: lie, truth) cells an alpha value was calculated. In four of the eight instances (all at the second rating point) the Cronbach's alpha was exactly 1.00. The remaining four alphas were above 0.90, indicating high overall reliability. Ratings were collapsed across communication channel, resulting in a 2 (consistency) x 2 (rating point) x 2 (veracity) design.

These consistency ratings determined which items in the stimulus set were relatively consistent and inconsistent. Having split the stimuli in this way, the aim was to determine whether the change in the PTJ over successive ratings differed according to the relative consistency of the speakers. The change in the PTJ between (i) the first and the second response and (ii) the first and the third response were calculated for the data collected in the previous studies that had access to the full behavioural repertoire. Namely, these were (i) both conditions of Experiment 1 (n = 83), (ii) the ratings from the full-length clips presented in Experiment 3 (n = 49), and (iii) ratings collected by Masip et al. (2009) from participants who viewed this stimulus set (referred to as the A1 video set in their publication) (n = 14). The change in the PTJ served as the dependent variable of the current analysis. All truth responses were coded as 1 and all lie responses were coded as 0, and the change in the PTJ was calculated as the judgment at the second (or third) response minus the judgment at the first response. As such, a shift from a truth (1) judgment after the first response to a lie judgment (0) after the second response was coded as (0 – 1 =) -1, a shift from a lie to a truth judgment was coded as (1 – 0 =) +1, and no change in judgment was coded as (0 – 0 = 1 – 1 =) 0. Thus a negative change value indicates a shift towards a lie response and a positive change value indicates a shift towards a truth response.

Results and Discussion

Consistency in these studies was a diagnostic indicator of deceit. Across the studies reported thus far it was generally the case that consistency was able to predict the decline in the PTJ.

Diagnosticity. A 2 (veracity: lie or truth) x 2 (time of rating: t_2 or t_3) within subjects ANOVA determined whether inconsistency ratings were significantly different between truths and deceptions. Such a difference would indicate that consistency is a diagnostic cue to deception. The main effect of veracity was found, F (1, 48) = 5.24, p = .026, η^2 = 0.10. Truths were rated as being more consistent (M = 4.50, SD = 0.74 n=49) than lies (M = 4.71, SD = 0.70). If raters make use of consistency, this might be considered adaptive insofar as it has an objectively valid basis.

There was also an effect of time, F (1, 48) = 9.64, p = .003, η^2 = 0.17, indicating higher perceived inconsistency when considering only the first and second response (t_2: M = 4.67, SD = 0.65) than when considering the inconsistency all three responses (t_3: M = 4.54, SD = 0.68). There was no significant interaction between the time of rating and the veracity of the statements, F (1, 48) = 0.05, p = .831, η^2 < 0.01.

Surprisingly, in contrast to prior research (Granhag & Strömwall, 1999, 2001b, 2002; Strömwall et al., 2003), inconsistency was an objectively valid indicator of deception (see Strömwall & Granhag, 2007, for evidence of consistency as a predictive cue of children's lies). This can be taken as an indication that a consistency heuristic (Granhag & Strömwall, 1999) is an objectively valid heuristic to employ, at least for the current stimulus set.

Consistency use. To determine whether raters were making use of consistency in their judgments, a set of 2 (consistency: low or high, within) x 2 (consistency up until rating point: second or third response, within) x 2 (veracity: lie or truth, within) ANOVAs were conducted on the change in the PTJ. Calculation of this variable is described in the Design section. Three separate analyses were conducted to determine whether consistency could account for the shift in the PTJ in (i) Experiment 1, (ii) the participants who viewed the full-length video clips from Experiment 3, and (iii) those who viewed the same stimulus set, video A1, in the original Masip et al. (2009) publication.

The first ANOVA, conducted on the ratings from Experiment 1, found a main effect of consistency, F (1, 82) = 9.12, p = .003, η^2 = 0.10 (Figure 6). Low consistency items showed a greater decrease in the PTJ (M = -.10, SD = .17) than high consistency items (M = -.03, SD = .14). In addition, a Consistency x Rating Point interaction was found, F (1, 82) = 18.93, p < .001, η^2 = 0.19. The change in the PTJ between t_1 and t_2 was significantly greater for low consistency statements (M = -.12, SD = .17) than for high consistency statements (M = -.001, SD = .15), as indicated by a Bonferroni-correct t-test, t (82) = -4.83, p < .001, d = -0.74. The change in the PTJ between t_1 and t_3, however, was not significantly different for high (M = -.08, SD = .20) versus low consistency items (M = -.06, SD = .15), t (83) = -0.58, p = .556, d = -0.11. There was no significant main effect of veracity, F (1, 82) = 0.05, p = .827, η^2 < 0.01.

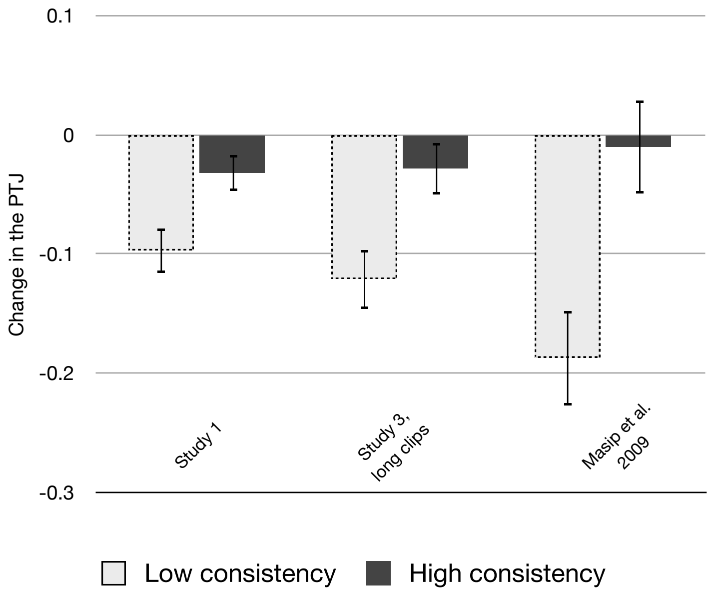

Figure 6. The change in the PTJ (overall) for high versus low consistency items, separated by study. Whiskers denote standard error.

The second ANOVA, conducted on the change in the PTJ in the ratings made of the full-length clips presented in Experiment 3, similarly found a main effect of consistency in the predicted direction, F (1, 48) = 8.39, p = .006, η^2 = 0.15, see Figure 6. There were no other significant main effects or interactions, including no effect of veracity, F (1, 48) = 0.03, p = .871, η^2 < 0.01.

The third ANOVA conducted on the change in the PTJ made by raters from the Masip et al. (2009) study again found a main effect of consistency in the predicted direction, F (1, 13) = 7.35, p = .018, η^2 = 0.36. There was also a main effect of veracity, F (1, 13) = 7.40, p = .018, η^2 = 0.36: judgments of deceptive responses showed a significantly greater decrease in the PTJ (M = -.16, SD = .12) than truthful responses (M = -.04, SD = .12). That is, across the responses raters were becoming more accurate, as reported by Masip et al. (2009).

Across the three studies conducted, consistency was a significant predictor of the change in the PTJ. Despite the use of a diagnostic cue, consistency, the decrease in the PTJ was not greater for lies than for truths with the exception of the 14 participants recorded in the Masip et al. (2006) study. This may suggest other cues aside from consistency may be factored into the truth judgment.

General Discussion: Chapter 4

In making lie-truth judgments, raters tend to show a bias towards believing (Bond & DePaulo, 2006; Zuckerman et al., 1979). Although there is relatively little research into the cognitive processes underlying lie detection (Lane & Vieira, 2012; Miller & Stiff, 1993; Reinhard & Sporer, 2008; Vrij & Granhag, 2012), a dual-process HAM approach has so far seen favour. Examining the conditions under which the truth bias is seen to be present or absent offers a means of testing the processes involved in leading to a truth bias. In one line of research, it is seen as arising form a default-interventionist system that initially is biased towards believing but is later interrupted by a more effortful evaluation (Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993; Masip et al., 2006; Masip, Garrido, et al., 2009; Masip et al., 2010). Here I tested the claims of these HAMs.

First, I considered whether the shift in judgments is a result of a shift in behaviour. No support was found for this account (Experiment 1). I then considered a more stringent test of the HAM: the amount of processing time should predict the degree of biased responding. Across three experiments I showed both experimentally and observationally that this was not the case. I also considered whether a class of

HAMs that claims an early selection of processing routes might account for the decline. However, no support was found for this class of model. Finally, an alternative to the HAM was sought. It was shown that consistency ratings could explain the decrease in the truth bias in the data presented here as well as in a prior study (Masip et al., 2009).

This section begins with a discussion of my findings in relation to past research and in relation to HAMs. I then consider my new interpretation of the phenomenon under a consistency explanation.

Heuristic-Analytical Accounts of the Truth Bias

In three experiments, the findings of Masip et al. (2006, 2009) were replicated, showing a decline of PJT across successive ratings. The predictions made by the HAM were not supported: the processing time up until the point of rating in Experiments 1 to 3 was unable to predict either PJT or accuracy, as default-interventionist and parallel-competition HAMs would predict (e.g., Evans, 2007; Gilbert, 1991; Masip, Garrido, et al., 2009). This was true whether viewing time was manipulated experimentally or tested statistically. There was also no support for the claim that participants may have shifted to analytical reasoning before the time of the first judgment: neither the PJT nor accuracy in response to the first portion of the statement could be predicted by viewing duration.

This lack of predictive ability of processing time fits with research showing a decline in the PTJ across a speaker's statement at much larger time scales, from 1 month to 5 months later (Anderson et al., 2002). It is also somewhat consistent with findings showing that with an equal viewing duration, those who made additional judgments between responses were more accurate compared to those who made a single judgment after viewing all three responses of a speaker (Granhag & Strömwall, 2001a). Whilst there was no increase in accuracy over successive ratings, the findings presented here also indicate that it is the act of making multiple judgments that causes changes in responding, not the amount of processing time available.

Perhaps more difficult for HAMs, there was a decline in PJT even when there was no initial truth bias, as was the case when only visual information was present in Experiment 2. Thus the decline in PJT was not the result of a shift away from biased responding that would be expected to result from a heuristic processing strategy.

A final class of HAMs, pre-emptive conflict resolution models, have also been considered in past research, albeit with a focus on accuracy rather than judgmental bias (Reinhard, 2010; Reinhard & Sporer, 2008, 2010). These models, however, also found no support in accounting for the truth bias: the types of information available did not result in the choice of heuristic or analytical processing from the outset (Chaiken, 1987; Gilbert & Krull, 1988; Reinhard & Sporer, 2008). These findings are consistent with research showing that people do not give more time to processing conclusions that are relatively difficult to believe (Ball, Wade & Quayle, 2006; Evans, 2007; Thompson, Striemer, Reikoff, Gunter & Campbell, 2003), as would be predicted by pre-emptive conflict resolution models: believability should be a readily accessible cue as to whether to select heuristic or analytical processing (Evans, 2007).

Although this class of HAMs predict a truth bias when only visual information is present and no (or a reduced) bias when verbal information is provided, there was no such effect. In fact, the means suggested, if anything, the reverse was true. This is perhaps surprising, because verbal content and speech cues are often thought to contain more diagnostic information than visual channels (Burgoon, Stoner, Bonito & Dunbar, 2003). My findings are not atypical, though: Bond and DePaulo (2006) conducted a meta-analysis on 50 studies and found (in both across-study and within-study comparisons) that visual only statements are rated as less truthful than statements that have access to audio only or full audio-visual channels.

There were two exceptions where a HAM could potentially have some explanatory power. First, in the audio-only condition of Experiment 2, the viewing time of the first response made could predict accuracy: with longer viewing durations there was greater accuracy. Of the four experiments presented, this was the only condition where visual behaviours were absent. One possible explanation of the finding is that speech cues are necessarily serial and take time to be fully presented. Visual behaviours can be produced in parallel. The increased accuracy with longer viewing times may reflect that fact that speech cues take time to become available.

This explanation requires that when both audio and video information is available accuracy increases with longer viewing times of the first statement. In Experiments 1 and 3 this was not found. One possible explanation of this is that video cues take primacy over audio cues. Indeed, this has been shown in a number of studies (Bauchner, Brandt & Miller, 1977; Chaiken & Eagly, 1983; Hocking, Bauchner, Kaminski & Miller, 1979; Maier & Thurber, 1968; Miller & Stiff, 1993; Nisbett & Ross, 1980). Because visual behaviours are presented rapidly and in parallel with other behaviours, the increase in accuracy over time may only be seen when these behaviours are absent, as in the audio-only condition of Experiment 2.

However, it is not possible to align these findings with a HAM. The model makes the claim that a shift between the types of processing modes leads to more reasoned and thus less biased responding. As a result of this shift, accuracy could potentially increase, but it need not (Chaiken et al., 1989). Thus a change in bias is a necessary condition whereas a change in accuracy is not. I found that the amount of processing time was unable to predict a change in bias in the audio-only condition.

Similarly, in Experiment 3 an increase in accuracy was observed for the thin-slice 8 s clips between the first and second response, but not between the second and third. No change in accuracy was seen for the long clips. This could suggest that between 8 s (the point of the first rating in the short clip condition) and 16 s (the point of the second rating) there was a change in processing style which could not have been observed in the longer clips that averaged approximately 30 s in length. However, as with the audio-only condition of Experiment 2, the change in accuracy was not accompanied by a change in bias. Again, it is difficult to align this with a HAM for the same reasons as those above. An alternative explanation was suggested in the discussion of Experiment 3 that took a similar tack to the explanation of the audio-only effect of Experiment 2. Accuracy in the short clips was initially very low, near chance. There is little information to be gathered after 8 s. After 16 s, more information became available, and accuracy increased to around 60%, where it levelled off. Accuracy in the long clips condition was maintained at around 60%. Thus the apparent increase in accuracy may simply reflect the lack of information available after the first thin slice of behaviour.

Thus far I have argued for a non-HAM interpretation of the findings. However, I do not attempt to claim that a heuristic-analytic model has no explanatory power in describing lie-truth judgments. In a number of experiments Reinhard and Sporer (2008, 2010) have shown the model does well at predicting when more or less accurate judgments are likely to be made. Accuracy was not the focus of the current research, but it is noteworthy that in Experiment 2 and 3 I found some evidence of accuracy differences.

It is also worth considering that I have tested the predictions made by various HAMs using a single phenomenon showing a declining truth bias across a speaker's statement (Masip et al., 2006; Masip, Garrido, et al., 2009). The pre-emptive conflict resolution model makes clear the temporal course of the two processes: one or the other process is selected at the outset. But the remaining two classes of model, the default-interventionist and parallel-competition models, are less clear on the temporal course. With longer processing durations an analytic process should be employed, but a number has never been put onto the boundary at which heuristic processing stops.

In the experiments conducted in this chapter, I have examined the judgment process over the course of minutes. However, it may be that the switch in processing occurs on the order of seconds. For example, Gilbert and colleagues (1990, 1993; Gilbert, 1991) propose a strong default-interventionist model. Their 'Spinozan mind' account claims the default process operates during the process of comprehension (i.e. as new behaviours are being presented) and that the intervening analytical process begins soon after. They show these switching effects between 0.75 s (Gilbert et al., 1990, Study 1) and 5 s (Gilbert et al., 1990, Study 2). In my experiments here, the shortest response made by any speaker was 8 s, which was also the selected length of the clip durations for the thin-slice condition of Experiment 3. A HAM may still be tenable at this more fine-grained time scale. This is examined experimentally in the next chapter.

The Role of Behaviour in the Truth Bias

To what degree do the behaviours of the speaker account for my findings? Experiment 1 sought to determine whether the truth bias could be explained simply as a reflection of the behaviours that are displayed by truth-tellers and that liars attempt to convey in their self-presentation. It was shown that regardless of whether a speaker's initial response or final response was rated first, there was a truth bias. Across multiple judgments the bias was seen to decline. This was taken as support for an independent cognitive component to the truth bias.

Yet Experiment 4 found that inconsistency in the speaker's response, a diagnostic indicator of their dishonesty, was able to account for the decline in the bias over time. Thus the phenomenon could be explained with recourse to the consistent or inconsistent behaviours displayed across the course of their statement. That is, there was a greater decline in the truth bias when the behaviours were inconsistent.

Consider also that Experiment 2 found an increase in accuracy with longer first responses of the speaker when given only audio information. One possible account of the finding suggested in the discussion of Experiment 2 was that the serial nature of speech, necessarily a temporally-extended cue, may account for the increasing accuracy: over time, more information became available to the rater. Similarly, in Experiment 3 accuracy increased for short clips between viewing 8 and 16 s, suggested to be due to the short amount of information available after 8 s. Again, there appears to be an exogenous influence on the judgment outcome.

Thus it can be seen that although there is a role for an independent cognitive component to the truth bias (as shown in Experiment 1), the behaviours of the speaker are incorporated into the judgment. Precisely how this interaction between environmental cues produced by the speaker and cognitive influences that the rater brings to the task will be an important question, a theme that runs through this thesis.

The Step-by-Step Account: Consistency

In the first experiment it became apparent that the decline in the truth bias was not attributable to processing time but rather to the number of judgments that had been made. This finding was replicated in Experiments 2 and 3. These findings are consistent with other research showing how the number of judgments made has a causal effect on subsequent judgments (Granhag & Strömwall, 2000b, 2001a; Hogarth & Einhorn, 1992). They are also consistent with research showing that even at much larger time scales, form 1 to 5 months later, there is a decline in bias with subsequent ratings of the speaker's statement (Anderson et al., 2002).

Because multiple judgments were made of a speaker, raters had the opportunity to make comparisons between statements. Granhag and Strömwall (2000b) found that when rating multiple responses from one speaker, 60% of the raters used consistency as a means of determining if the speaker was lying. It has been shown that consistency appears to have little diagnostic value as a means of detecting deception (Granhag & Strömwall, 2001a; Strömwall et al., 2003), although both police officers (Greuel, 1992), and laypersons (Granhag & Strömwall, 2000b, 2001a) agree that consistency implies truth telling and inconsistency implies deceit.

Indeed, in my studies they would be well placed to hold that belief: consistency was in fact predictive of deception with the current stimulus set. The differences between my studies and those of Granhag and colleagues may be attributable to the differences in the nature of the lies told or some other difference that can be observed in the behaviours of the speaker. I cannot comment on why in my experiments consistency was a valid indicator of deception other than to note that it was. It is interesting to ask whether raters would make use of consistency information even if it were not diagnostic. The research of Granhag and colleagues discussed above suggests that they would: raters in their studies appear to make heavy use of consistency despite its non-diagnosticity.

What is perhaps surprising is that despite being diagnostic in my studies, accuracy did not increase with between the first and second rating, where the option for judging consistency between responses became available. There are at least two possible explanations. First, it may be that raters were not only making use of consistency information but also other cues, and it was these other cues that resulted in the lack of an increase in accuracy. In a recent study (Bond et al., 2013), participants were given a 100% diagnostic cue to deception: the incentives the person had for lying or telling the truth. When only this information was available, accuracy was near perfect, around 97%. In one of their experiments, participants were also given the audio-visual behaviours from the speaker's statement. Because raters were making use of a 100% reliably diagnostic cue, there would be no need to integrate other cues into the judgment. Nonetheless, they did, and on doing so their accuracy suffered markedly, dropping to around 75% (Bond et al., 2013, Study 3). In contrast to one-reason decision-making accounts (Gigerenzer & Goldstein, 1996, 1999; Gigerenzer et al., 2008; Kyonka & Church, 2005), this may suggest other information is also being taken into account during the decision. Indeed, there is evidence that even when simple heuristics are being used other information is also brought to mind (Khader et al., 2011; Platzer & Bröder, 2012). This is considered further in Chapter 7.

There is another interpretation. Consistency was only measured between responses. Raters may have also been making use of the internal consistency of a single response, which may also have been diagnostic of deception. This is speculative, but a plausible alternative explanation. Distinguishing between the multiple-cue-use hypothesis above and the continued-consistency-use hypothesis here is not possible with the current data.

Conclusions: Chapter 4

In the first three experiments I have shown that, whilst judgments of truth are often high initially, they decline over successive ratings of a speaker's statement. Contrary to an often-cited explanation of the truth bias, no support was found for a HAM account of either the initial truth bias or the decline in the PJT. It should be stressed that no claim is made that the findings here discredit heuristic models, or even that heuristics are not used when making lie-truth judgments. The findings indicate that an initial tendency towards believing a statement as truthful is not driven by heuristic processing, and that the shift away from truth responding as the speaker continues to deliver their statement does not reflect a shift towards an analytical processing strategy.

It was shown that the decline in the bias was attributable to the act of making multiple judgments. Further research showed that the use of a relatively simple heuristic, the consistency heuristic, could account for the decline: statements that showed less consistency between the first and second response were less likely to be judged as truthful by the point of the second judgment. Instead, it is suggested a single process account can better account for the findings. What this single process system may look like is considered in the next chapter.

Chapter 5: The Adaptive Decision Maker

To date there has been little exploration of the cognitive processes underlying lie detection. This is surprising for two reasons. First, if the largely applied field of lie detection is to boost accuracy for law enforcement officials, it will need to consider how they make their decisions and why they so often go wrong. Second, lie detection is a real-world socially oriented decision making task that allows researchers to consider how contextual and social information influence the judgment process while maintaining experimental control. Those who have begun to consider the lie detection process have taken a HAM position. The previous chapter took that position as its starting point.

A decline in the truth bias across time has been taken as evidence of a default-interventionist HAM, where heuristic processing is interrupted upon by a more analytical form of processing after a given amount of time. My research indicates there is little support for the HAM, and instead I showed that, although not the full picture, perceived consistency in the speaker's responses could better explain the decrease in truth judgments over time.

However, it could be argued that the time scale of minutes is too coarse to explore the shift in processing (although see the thin slice condition of Experiment 3). This chapter begins by considering a model at a finer temporal resolution that has influenced thinking in philosophy (Burge, 1993; Millikan, 1987) as well as in the psychological domain, including hypothetical reasoning (Fitzsimons & Shiv, 2001;

Moore et al., 2012), text comprehension (Hasson et al., 2005; Prentice et al., 1997; Schul et al., 2004), persuasion (Green & Brock, 2000; Sperber et al., 2010) religious belief (Pennycook et al., 2012), and lie detection (Colwell et al., 2012; Levine et al., 1999; Millar & Millar, 1997). The *Spinozan mind hypothesis* (Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993) proposes two stages of processing. In order to comprehend, the hypothesis claims it is necessary to believe the information to be true. In a second stage preceding comprehension, a more critical deliberation can lead to a reclassifying of belief from accepting the information is true to believing it is false. Because comprehension is online, which is to say that we comprehend each new piece of information as and when it becomes available (e.g. Heuttig et al., 2011; Spivey et al., 2005; Spivey, 2007), the effects of initial belief due to comprehension and the subsequent revision of information are proposed to take place within the initial few seconds of receiving information (Gilbert et al., 1990).

Gilbert and colleagues (Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993) contrasted the Spinozan mind with a Cartesian mind. The Cartesian approach claims there is no initial automatic acceptance of information. Instead, there is an initial period of uncertainty during comprehension, after which a decision can be made as to whether the statement is believed to be true or false. As such, it is a single processing model, in contrast to the dual-process model of the Spinozan mind. To distinguish these two accounts, it is necessary to examine the decision process during comprehension. It is only during the initial stages that the Spinozan and Cartesian mind accounts differ. The researchers interrupted the formation of judgments early in the comprehension process. When participants were tested for their memory of the statements, they were more likely to believe a statement was true than to believe it was false. This was taken as evidence that during comprehension there is a bias towards believing information to be true. This paradigm underlies the direct support that the Spinozan mind hypothesis has received (Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993).

However, the evidence is compromised by the structure of the task. An assumption is implicit in the two-alternative forced choice task: if there is uncertainty, there should be an equal number of responses for each of the options. That is, participants should be no more likely to use one response than another. The naïve Cartesian may be expected to guess randomly. The data do not support this, and so the naïve Cartesian can be rejected. But the random guessing assumption is violated in many experimental settings. Participants arrive at the laboratory with a history of experience. Unsurprisingly, people tell the truth far more often than they lie (Caspi & Gorsky, 2006; DePaulo & Kashy, 1998; DePaulo et al., 1996; George & Robb, 2008; Hancock et al., 2004; see also Cole, 2001). An informed Cartesian model would predict that even during the early moments of processing, when there is initial uncertainty, if forced into making a judgment we rely on our prior knowledge or information about the current context.

Models of language comprehension and belief formation have taken a similar view (see Johnson-Laird, 1983; Richter, Schroeder & Wohrmann, 2009; Schwarz, Sanna, Skurnik & Yoon, 2007; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). Even in situations where effort has been taken to remove the possibility of using prior knowledge, such as word-learning studies using nonsense words (Gilbert et al., 1990), participants may make use of the meaningfulness (Beukeboom et al., 2010; Hasson et al., 2005; Wegner et al., 1985) and ambiguity (Anderson et al., 2009; Fraundorf et al., 2013; Glenberg et al., 1999; Mayo et al., 2004; Skowronski & Carlston, 1989) of the definition or make use of contextual knowledge: e.g. in teacher-student environments like the language-learning task, teachers rarely provide inaccurate information (see Fiedler et al., 1996; Grice, 1975; Swann et al., 1982).

By removing the necessity to generate a response, a Cartesian system need not make use of prior knowledge to make an informed guess; instead, they can explicitly indicate their judgmental uncertainty. Experiment 5 shows how, during the comprehension of new information, removing the necessity to make a judgment reduces the truth bias. Raters show an initially high degree of uncertainty that over time declines in favour of making a judgment, as would be expected of a Cartesian mind. It is suggested the use of prior knowledge may be used when forced into a judgment despite being undecided.

Past knowledge may take the form of an availability heuristic (O'Sullivan et al., 1988): there are more instances of truthful interactions in memory than instances of deceptions. Experiment 6 explores this heuristic further by explicitly manipulating the perceived base rate of honesty. During the comprehension of information, raters make use of their expectations of honesty to guide their judgments. After comprehension and at the point of making a lie-truth response, there is an evident truth bias. Base rate knowledge partially influences the truth bias insofar as it can amplify the bias, but expecting most speakers to lie does not result in a reduced truth bias at the point of making the response. This finding is consistent with the notion of an early use of more general prior knowledge in the absence of information and when forced into a judgment, followed by a switch towards making use of other decision-making strategies when the information is present but ambiguous.

Because the bias seems to result from a forced choice rather than from the act of online comprehension, Experiment 7 considers whether the truth bias can be reduced after comprehension by allowing raters to indicate their uncertainty. Surprisingly, the ability to explicitly withhold judgment results in an *increased* truth bias. Upon exploration it can be seen that the increase in the bias is artefactual insofar as a relatively large proportion of the lie judgments in a forced lie-truth choice become unsure judgments when allowed to indicate uncertainty. This effect is somewhat replicated in Experiment 8, a replication of Gilbert et al. (1990, Study 1), although it should be noted that on exploration of the data the design of the study was not well suited to test the Spinozan or Cartesian account.

The findings are considered evidence of the adaptive, flexible and situation-dependent strategies that raters employ. When information has been gathered from the environment and when this information is not sufficient to make a distinction, raters rely on their past experience. This experience comes in the form of metacognitive certainty: raters are more confident in judging truths than in judging lies (Anderson et al., 2002; DePaulo et al., 1997; Anderson, 1999, cited by DePaulo & Morris, 2004; see also Levine et al., 1999). In a forced judgment, uncertainty in spite of sufficient available information can be taken as indication of the possibility of the statement being a lie. This form of uncertainty is internal to the decision maker. When there is a lack of information in the environment (Experiment 5), uncertainty can be considered as due to external factors. External uncertainty can no longer be used as a useful heuristic towards guessing the statement as deceptive. In the lack of available evidence in the environment, raters can rely on other past knowledge in the form of experience with the world – people usually tell the truth – a heuristic that has objective validity (Caspi & Gorsky, 2006; Clark & Clark, 1977; DePaulo et al., 1996; Fan et al., 1995; George & Robb, 2008; Hancock et al., 2004; Levine, Kim & Hamel, 2010; O'Sullivan, 2003; O'Sullivan et al., 1988; Serota, Levine & Boster, 2010; van Swol et al., 2012).

Experiment 5: A Cartesian Mind - Online Comprehension

To compare the Spinozan mind with the modified Cartesian account, an affirmation-negation task was derived from an experiment Gilbert et al. (1990, Study 2) carried out to support the Spinozan view. Participants judged whether another person was being honest or deceptive. Two key changes were made to Gilbert et al.'s (1990) methodology: the number of responses participants could use, and how often they were sampled. Gilbert et al. (1990) obtained a single truth or lie judgment from raters by interrupting their decision process and sampling from that time point, which reveals little of how that decision process unfolded up until that point (Spivey, 2007; Spivey & Dale, 2006; Thelen & Smith, 1994; see also Allen & Ebbesen, 1981; Busemeyer & Townsend, 1993; Carlston & Skowronski, 1986).

The Spinozan account rests upon the time course of the bias. To give a window onto the time course of this decision process, in the present study raters made continuous judgments across the full duration of a speaker's statement. To foreshadow the results, this fine-grained measurement was able to replicate the Spinozan effect found in more coarse measures. That is, raters were truth biased even in the earliest moments of processing when asked to give a lie-truth response, replicating the findings of Gilbert and colleagues (1990, 1993).

This could be because they are Spinozan raters and must automatically believe the statement in order to comprehend it; or, they are Cartesian raters, utilizing their past knowledge when forced to make a judgment. By giving the option to indicate their uncertainty, it is hypothesised that raters will appear Cartesian: that is, they will show a reduced or no bias because the rater is relieved of the necessity to lean primarily on their pre-existing knowledge in order to select one of the experimenter's presented options of either belief or disbelief. Put another way, a Spinozan form of responding is anticipated in the forced binary choice condition where participants have no option but to make a decision, even if they are entirely unsure. But the apparent automaticity claimed by the Spinozan view is expected to disappear when raters can abstain from judgment.

Note that if the truth bias is a reflection of the system's requirement to accept information (i.e. the Spinozan view), the addition of the unsure option should have little influence on (excessive truth) responses because such automatic processes cannot be overcome by deliberate attempts to suppress them (Neely, 1977; Shiffrin & Dumais, 1981). If raters are indeed Spinozan one would predict either no use or, less likely, random use of the unsure button. In both cases, a tendency towards believing over disbelieving would still be observed after removing the unsure responses.

Method

Materials. To explore truth-lie judgments, a stimulus set from a prior study was used (Street et al., 2011, April). Twenty-two participants were approached by a junior researcher posing as an assistant to a documentary director. The researcher claimed to be looking for volunteers to interview on camera about their holiday experiences. He determined a small set of countries participants had and had not visited. The researcher said he was short of time and had been unable to find anyone who had visited some countries. So as a favour he asked the participants to tell the director they had spent time in one of the countries they had never visited as well as giving a monologue about a country they had truthfully visited. The researcher stressed the director would believe they were telling the truth throughout. If participants agreed, they were taken to a live filming studio where they were left alone with another experimenter posing as the director. The director stressed the importance of filming genuine accounts of their experiences in both countries. Participants signed a waiver to this extent stating that they would be entirely truthful in both their deliveries. The director was blind to which of the two statements would be deceptive, with the order of statements counterbalanced by the junior experimenter. Participants were positioned in front of three visible cameras such that their head and torso were visible. They were asked to speak for approximately 30 s in response to the question, “When you arrived in [country name], what was your first impression of the people there?”. After delivering both statements, participants were fully debriefed and given the opportunity to retrospectively withdraw their consent. No participants chose to withdraw their consent. Participants volunteered with no monetary compensation.

Two speakers were excluded from the final stimulus set because they admitted to the director they were about to lie. Two further participants were selected as practice stimuli. This left 18 speakers each delivering a lie and truth, resulting in 36 videos. From these, two stimulus sets were created such that the 18 speakers appeared only once in each set and that each contained nine lies and nine truths. Statements ranged between 10 s and 91 s. Truthful statements lasted 32.86 s (SD = 10.79) on average, while deceptive statements lasted 32.72 s (SD = 24.83). Although the average durations are similar, the variance is larger in the deceptive statements which may be an indication of certain individuals being unsure what to say while others being able to deliver lengthier lies (see Levine, 2010, for consideration of individual differences in the ability to lie). It should be noted that although we asked participants to speak for approximately 30 s they were not constrained to talk for this period. The duration of a statement may itself be a valid and/or perceived cue to deception.

These statements make up the Bloomsbury Deception Set (Street et al., 2011, April), and are used throughout this chapter. Videos were presented and responses collected throughout this chapter using MATLAB software (The Mathworks inc., 2000: Natick, MA) and the Psychophysics Toolbox (Brainard, 1997).

Participants. Forty-six rater participants took part and were compensated £3 for their time. One participant was excluded because their responses indicated that they did not understand the instruction. This left 25 females and 20 males with a mean age of 26.0 years (*SD* = 7.7 years, range = 18 to 54 years).

Procedure. The written instruction explained that each speaker would lie or tell the truth about people they claimed to have met in a foreign country. They were also instructed they should continuously rate the statement as it was being delivered. Throughout each statement, participants indicated moment by moment whether they currently believed or disbelieved the speaker by holding down either the left or right arrow keys. Participants were instructed they must respond at the onset of the video. In this way, a continuous measure of the lie-truth judgment trajectory was captured. If both buttons were pressed simultaneously participants were instructed to release both keys and begin pressing with only one key. In the event of two buttons being pressed together, the system recorded the first of the two keys that the participant pressed. There was no on-screen feedback about which key they were holding.

This allowed us to sample the unfolding judgment rather than interrupting the process as Gilbert and colleagues did. The continuous response was binned into five proportional time points afterwards for ease of analysis. Participants in the lie-truth (LT) condition (*n* = 23) indicated throughout the video whether the speaker was lying or telling the truth. Participants in the lie-truth-unsure (LTU) condition (*n* = 22) were given the additional option of indicating their uncertainty.

At the end of each video participants in both conditions made a binary lie-truth response typical of lie detection experiments. That is, regardless of experimental manipulation, all participants at the point of the final judgment made a forced binary choice. The main aim of this study was to determine how the belief trajectory developed over time. For this reason, the end of statement responses were not of primary interest and so are not discussed further. There were two practice videos, after which the instruction was presented again and the remaining 18 experimental trials given.

The selected video set, the position of the lie and truth response options on the screen, and whether participants were asked if the last speaker was 'lying or telling the truth' versus 'telling the truth or lying' were fully counterbalanced between participants and conditions.

Design. The independent variables were the in-trial response conditions (LT or LTU), the veracity of the speakers' statements and the proportional time that had elapsed. For ease of analysis, the proportional time was binned into 5 discrete time points. The dependent variable was the proportion of truth judgments, resulting in a 2 (response condition: LT or LTU, between subjects) x 2 (speaker veracity: lie or truth, within subjects) x 5 (time point, within subjects) mixed design.

Results

Over the course of the statement raters in the LTU condition reported they were unsure on average between 35% (*SD* = 22.06) at the first binned time point and 17%

(*SD* = 12.16) at the last binned time point. In both conditions all participants made a forced binary choice at the end of the statement. During the early moments of processing there was a bias towards truth believing, replicating Gilbert et al. (1990), but only amongst those forced into making a binary lie-truth decision. Those able to indicate their indecision showed no such bias, instead exhibiting a pattern of responding consistent with a Cartesian rater.

Belief trajectory. One could argue that a Spinozan rater may use the unsure button randomly. The presence of an additional response could artefactually decrease the proportion of truth judgments (PTJ). To prevent this, and to allow for comparisons of the belief trajectories between the LT and LTU conditions to be made, the PTJ of all the lie-truth responses was calculated. That is, in the LTU condition the unsure responses were discarded entirely. If they were kept, chance responding would be at 0.33 because responses could be expected to be equally distributed amongst the three options. By discarding the unsure responses and examining only the use of the lie and truth buttons in the LTU condition, an equal (unbiased) distribution between these two responses would be 0.5, the same as is the case of the LT condition. Thus a bias to responding true (i.e. a ‘truth bias’) would be demonstrated if the PTJ were greater than half the responses made, irrespective of experimental condition. All findings in this and the following section discuss precisely these data.

Since stimulus speakers provided spontaneous speech, strict durations for their statements could not be imposed. They varied from 10 to 91 s, with the average statement lasting 32.79 s (*SD* = 18.83). To compare across items the PTJ was binned into five equally spaced time points. In the following section the PTJ data are reanalysed exploring the first 2 to 10 s of the judgment period. It is during these early moments of processing that Gilbert et al. (1990) found a Spinozan truth bias. To anticipate those analyses, the findings exploring the first 10 s mirror those exploring the proportional time across the course of the statement.

Supporting the hypothesis, a 2 (response condition: LT or LTU, between) x 2 (speaker veracity: lie or truth, within) x 5 (time point: t_1 - t_5, within) mixed ANOVA was conducted on the PTJ. A main effect of response condition was found, F (1, 43) = 5.57, p = .023, η^2 = 0.12 (see Figure 7). Having the option to indicate indecision resulted in a smaller truth bias both during the early moments of processing and across the remainder of the trial, compared to raters who were forced into a binary response. This main effect was not moderated by time point or speaker veracity, all *p*s > .1.

There was a main effect of time, F (1.43, 61.56) = 3.81, p = .041, η^2 = 0.08, which interacted with veracity, F (2.16, 92.95) = 5.94, p = .003, η^2 = 0.12. Post-hoc Bonferroni-corrected *t*-tests indicated that the PTJ did not change significantly over time when rating deceptive statements, all *p*s > .1, but when rating truthful statements the PTJ increased from t_1 to t_2, t (44) = 3.60, p = .006, d = 0.33, and from t_2 to t_3, t (44) = 3.05, p = .046, d = 0.27, but did not increase further, all *p*s > .1. That is, the PTJ ran counter to the predictions of a decrease over time as would be made by a Spinozan account.

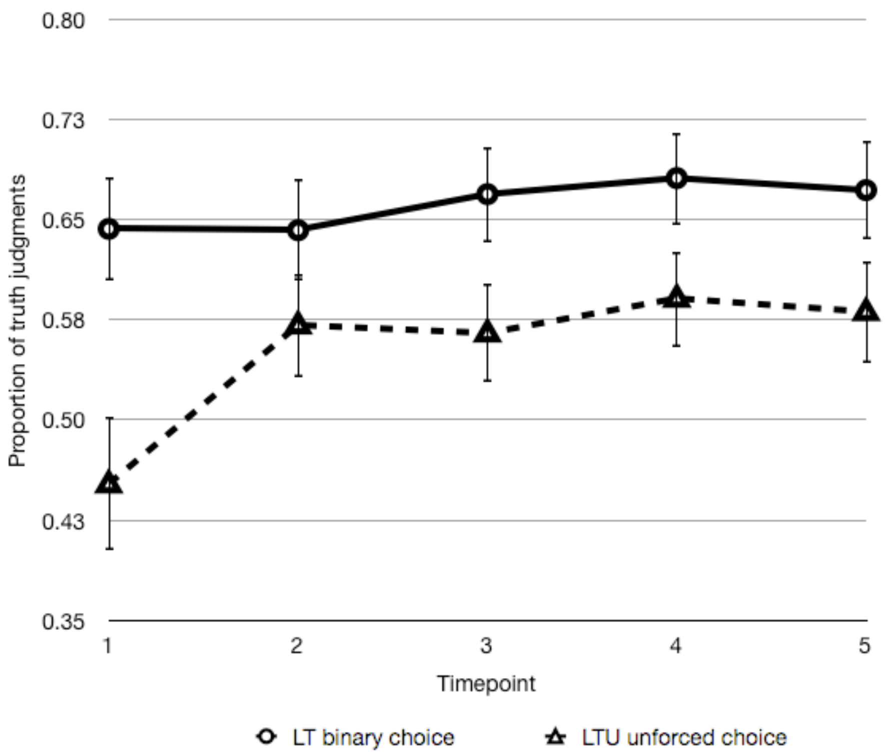

Figure 7. The proportion of truth judgments made in the LT and LTU conditions across the duration of the trial, split into five equal proportional time bins. Whiskers denote standard errors throughout the chapter. Note that the unsure responses from the LTU condition have been removed, and that chance PTJ is at 0.5 for both the LT and LTU conditions. Whiskers denote standard error.

The main effect of condition, taken as support for the hypothesis, could be attributable to increased accuracy rather than a reduced bias. Signal detection analyses examined the independent effects of accuracy and response bias. Because the difference between conditions is of primary interest, which did not interact with time, the PTJ was collapsed across time points. An independent samples *t*-test found no significant difference in A' between the LT and LTU conditions, t (32.99) = 1.77, p = .861, d = -0.03. One-sample *t*-tests determined A' was not above chance rate in either the LT or LTU conditions (all *p*s > .1). More importantly, there was a significant difference in B''_D between the LT and LTU conditions, t (32.57) = 2.34, p = .026, d = 0.70. One-sample *t*-tests confirmed LT raters (M = 0.35, SD = 0.23) were truth biased, t (22) = 7.38, p < .001, d = 1.54, whilst LTU raters (M = 0.12, SD = 0.41) were not, t (21) = 1.36, p = .189, d = 0.29. A Bayes factor using the JZS one-sample *t*-test equivalent determined that although the data were rather decisive for the presence of a truth bias when forced to judge (Bayes factor of around 50,000), there was negligible evidence in favour of a lack of a truth bias, with the data being only 1.6 times more likely under the hypothesis of a lack of response bias.

Consider that the Spinozan effect makes predictions based at the earliest moments of processing, a position discussed in more detail in the next section. Two further one-sample *t*-test equivalents were conducted on the first proportional time point. It was found that for the forced judgment condition the data were 10 times more likely under the hypothesis of a truth bias being present, but when not forced into judgment the data were 3 times more likely under the hypothesis of a lack of bias. This provides evidence that, during the early moments of the trial, there is a lack of bias if not forced into judgment. This is considered further in the next section.

Early moments of processing. As the average statement lasted 33 s, the first proportional time point reflects the decision process 6.6 s in for the average trial. Gilbert and colleagues demonstrate that judgments obtained as early as 10 s after initial presentation of a statement (8 s statement presentation plus 2 s delay before response) were truth biased. The data are re-analysed examining the first 10 s of the average trial by binning the data into 2 s windows. This also addresses the fact that statements varied in length.

Not all participants began responding at the point of video onset, despite instruction to do so. This significantly reduced the number of participants with full cells at the two-second period (n_{LTU} = 8, n_{LT} = 12). Nonetheless, the findings closely mirror those conducted above. A 2 (condition) x 2 (veracity) x 5 (time point: 2 s to 10 s, within subjects) mixed ANOVA with the PTJ as the dependent variable demonstrated a main effect of condition, F (1, 18) = 4.33, p = .052, η^2 = 0.19, with the PTJ higher in the LT (M = .60, SD = .31) than the LTU condition (M = .37, SD = .36). In addition, there was a Veracity x Time point interaction, F (1.44, 25.92) = 4.43, p = .033, η^2 = 0.20. Post-hoc Bonferonni-corrected analyses were all non-significant, potentially because of the small sample sizes, but the means indicated a decline in PTJ for deceptive statements from 2 s (M = .54, SD = .40) to 10 s (M = .46, SD = .18), whilst the PTJ for truthful statements increased from 2 s (M = .40, SD = .40) to 10 s (M = .56, SD = .31). There was no main effect of veracity, F (1, 18) = 0.16, p = .690, η^2 = 0.01.

Signal detection measures were all non-significant, all ps > .1, with the exception of an independent samples t-test that, although only marginally significance, found the LT condition (M = .24, SD = .54) was more truth biased than the LTU condition (M = -.32, SD = .75), t (18) = 1.95, p = .067, d = 0.86.

Cartesian responding. Thus far the use of the unsure response has been ignored in order to compare the LT and LTU conditions. Confirmatory support for a Cartesian mind can be found in the way in which raters use the unsure response option. A Cartesian account predicts a high proportion of uncertainty early on that declines over time. As noted in the previous section, not all participants began responding at the point of video onset despite instruction to do so. For that reason, the proportion of responses during the early moments of video onset does not sum to 1: some participants chose to make no response at all.

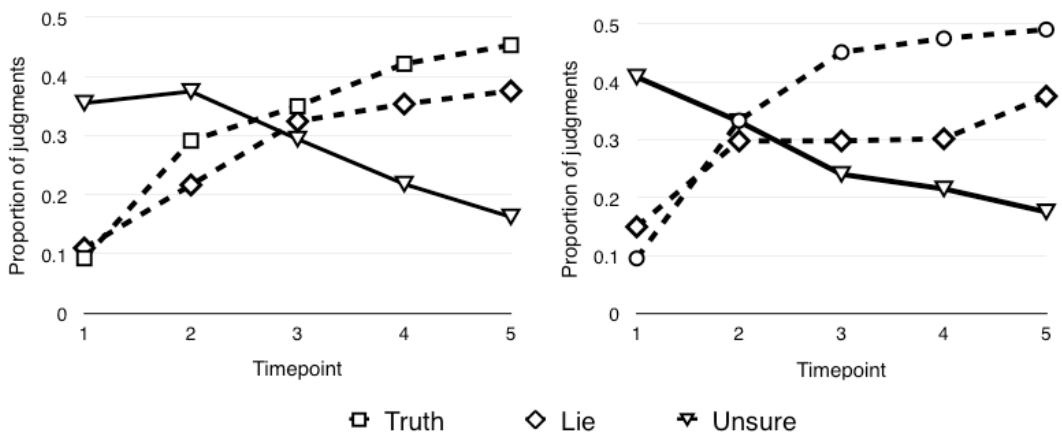

Figure 8. (a) The proportion of truth (square), lie (diamond) and unsure (inverted triangle) responses across the duration of truthful statements. Unsure responses are relatively high during the early stages of processing and decline over time in favour of lie and truth responses (left). (b) The proportion of truth, lie and unsure responses across the course of deceptive statements (right).

A one-way ANOVA conducted on the proportion of unsure responses across the five time points found a significant effect of time point, F (1.40, 29.41) = 11.92, p = .001, η^2 = 0.36 (Figure 8). Bonferroni-corrected t-tests found the proportion of unsure responses at time point 1 (M = .38, SD = .28) did not differ significantly from time point 2 (M = .35, SD = .23) or 3 (M = .27, SD = .15), both ps > .2.84, but was significantly greater than at time points 4 (M = .22, SD = .14) and 5 (M = .17, SD = .12), both ps < .036, ds = 0.72 and 0.98 respectively. Similarly, unsure responses at time point 2 were significantly greater than at points 4 and 5, both ps < .013, ds = 0.68

and 0.98, but not at time point 3, t (21) = 2.87, p = .087, d = 0.41. There were not significantly more unsure responses at time point 3 than at time point 4, t (21) = 3.00, p = .062, d = 0.35, but unsure ratings at time point 4 were significantly greater than at time point 5, t (21) = 5.33, p = .001, d = 0.38. Thus across the course of the statement the proportion of unsure responses was seen to decline, as would be predicted by a Cartesian but not a Spinozan account.

Discussion

Previous research has shown raters are biased towards accepting information is true in the earliest stages of comprehension (Gilbert et al., 1990; Gilbert et al., 1993). As Gilbert et al. (1990, p. 601) put it, ‘all ideas are accepted… *prior* to a rational analysis of their veracity, and that some ideas are subsequently *unaccepted*’ (italics in original). In contrast, a Cartesian account predicts a period of initial uncertainty that gives way to a judgment after comprehension. Given the differing time predictions of the two accounts, here the judgment process was tracked across its trajectory. First, the general Spinozan phenomenon was replicated: if forced into judgment raters were biased towards believing what others said, even during the early moments of processing. However, if they were able to indicate indecision (LTU condition), people acted like Cartesian raters, which was particularly noticeable during the earliest moments of comprehension. It would seem people do not merely believe what they are told, and can comprehend without having to automatically assign a belief value.

Models of judgment formation have often taken a Cartesian approach, either implicitly or explicitly assuming uncertainty rules at first until some threshold or condition is met (Allen & Ebbesen, 1981; Busemeyer & Townsend, 1993; Carlston & Skowronski, 1986; Ratcliff, 1978; Ratcliff, Schmiedek & McKoon, 2008; Roe, Busemeyer & Townsend, 2001; van Ravenzwaaij et al., 2012), but that even from the outset the judgment can be biased towards preferring one alternative (e.g. Roe et al., 2001; van Ravenzwaaij et al., 2012). Prior experience (Bransford & Johnson, 1973; see also DePaulo et al., 1996; Fan et al., 1995; O'Sullivan, 2003) and expectations (Kassin et al., 2005; Masip et al., 2005; McCornack & Parks, 1986, 1990) can serve to bias the direction of judgment from these early moments.

When unsure but required to make a judgment, the logical choice would be to select the option for which there is most evidence, whether this evidence is built up from prior experience or directly from the stimulus item. Assuming our stored knowledge of prior experiences with the world, for example, has at least some basis in reality (or is at least perceived to hold some validity), utilising this knowledge when there is no other information to work from will result in greater accuracy than random responding. Thus under conditions of uncertainty a rater can *appear as though* they are automatically accepting statements as true but may in fact be responding adaptively and appropriately given their uncertainty.

The findings from the Spinozan and Cartesian camps have been aligned by showing that when a Cartesian rater is forced to affirm or deny a belief they appear distinctly Spinozan, but when able to express their indecision they once again appear Cartesian. The mind is able to comprehend information before having accepted or denied it as the truth. Yet if pressed for a judgment before one has been reached, we are sufficiently flexible to be able to incorporate prior knowledge and experience into the judgment.

Experiment 6: Most People Tell the Truth – The Availability Heuristic

The current study considers what sorts of information may be used under uncertainty. If raters come to rely on their prior knowledge with similar situations, one likely possibility is that they will make use of base rate information: that is, how often people typically lie or tell the truth.

That the truth bias reflects the base rate is certainly not a new suggestion (Clark & Clark, 1977; DePaulo & Rosenthal, 1979; O'Sullivan et al., 1988), but to date there have been no direct test of the claim. There is some strong suggestive evidence in favour of the use of base-rate information. For example, there is a greater truth bias when rating those with who we are in close relationships (Argyle & Henderson, 1984; McCornack & Levine, 1990; McCornack & Parks, 1986; Stiff et al., 1992; Stiff et al., 1989; van Swol et al., 2012; see also Boyes & Fletcher, 2007; more on this in Chapter 6), people that we expect to tell us the truth (Gaertner & Dovidio, 2000; Voci, 2006), whilst police investigators, who tend to disbelieve others more often than not (Ask, Rebelius & Granhag, 2008; Meissner & Kassin, 2002), expect in general that their interviewees will lie to them (Kassin et al., 2005; Masip et al., 2005; Moston et al., 1992).

Making use of base-rate information is adaptive insofar as it increases accuracy (Kahneman & Tversky, 1973; Nisbett & Ross, 1980). For instance, people tell fewer lies to close relational partners compared to strangers (Cole, 2001; DePaulo & Kashy, 1998), and in general people tend to tell the truth more often than they lie (Caspi & Gorsky, 2006; Clark & Clark, 1977; DePaulo et al., 1996; Fan et al., 1995; O'Sullivan, 2003; O'Sullivan et al., 1988; van Swol et al., 2012).

Although early research has suggested such a neglect of base rate information in favour of using more individuating information (Kahneman & Tversky, 1973; Nisbett & Ross, 1980; Tversky & Kahneman, 1974), in recent years research has begun to question whether the base rate is taken into account, with research showing how it can influence decision making (Bar-Hillel, 1980; Barbey & Sloman, 2007; Ginossar & Trope, 1980; Koehler, 1996; although see Kahneman, 2003) and social judgments (Funder, 1996). That early research found no effect of base rate information may be a result of how that information was framed: reframing the information reduces neglect of the information (Hilton & Slugoski, 2001; Schwarz, Strack, Hilton & Naderer, 1991; Zukier & Pepitone, 1984). That is, base-rate information use is context-dependent (Kruglanski et al., 2007)

In the current study participants' beliefs about the base rates of lies and truths was explicitly manipulated to determine its effect on the truth bias both during the development of the judgment, while the statement was being delivered, as well as at the point of the final judgment, after the information from the environment had been obtained. It was predicted base-rate expectations would influence responding, particularly during the early moments of statement comprehension because it is here that participants will have little behavioural information from which to work from.

In addition, greater confidence was predicted in making 'truth' compared to 'lie' ratings when the majority of the statements were expected to be truthful, whilst the reverse was predicted when the majority of the statements were expected to be deceptive.

Method

Materials. The Bloomsbury Deception Set (Street et al., 2011, April) used in Experiment 5 was used here.

Participants. Ninety-nine first year psychology students participated as part of their undergraduate studies. Fifteen participants were excluded because they had taken part in a similar study for course credit at an earlier date. One participant was excluded because of hearing difficulties, leaving 83 participants (67 females) with a mean age of 18.7 years (*SD* = 0.70 years, range 17 to 21 years).

Procedure. Participants were tested in isolation. Instructions explained the speaker had provided a truthful and deceptive account about having visited a holiday destination, but that they would only see one of these accounts. As in Experiment 5, a continuous measure of the lie-truth judgment trajectory was captured. Each video was counted in from three to one before onset. After each video, participants were instructed to indicate whether they believed the speaker was lying or telling the truth by clicking with the mouse one of the two response options presented on screen. They also rated how confident they were in their response on a scale of 1 (not at all confident) to 10 (very confident).

Participants in the low truth expectancy condition (*n* = 26) were told only '20% of the speakers were telling the truth that they had visited the country, 80% of speakers lied. That is, most people lied.' The high truth expectancy condition (*n* = 27) was told 80% of speakers told the truth, whilst the equal split expectancy condition (*n* = 30) was told half the speakers lied and half told the truth. Note the base rate information is presented both as a normalised frequency, i.e. a percentage, as well as in a more accessible fashion, i.e. text explicitly explaining whether the majority of the speakers lied or told the truth, making the base rate information more readily comprehendible (see Barbey & Sloman, 2007; Gaissmaier, Straubinger & Funder, 2007).

In reality, all conditions saw an equal split of lies and truths. Assignment of video set, position of lie-truth buttons on the screen, and question order ('did the last person lie or tell the truth' versus 'did the last person tell the truth or lie?') were counterbalanced.

Design. Participants were randomly assigned to experimental conditions. The independent variables were the anticipated proportion of truths to lies, i.e. either 20%, 50% or 80% of truths were expected, and the veracity of the speaker's statement. The dependent variables were the PTJ, confidence and accuracy taken after each statement had ended. In addition, signal detection measures of bias and accuracy were taken to assess the effects of bias and accuracy independent of each other. This resulted in a 2 (veracity: lie or truth, within subjects) x 3 (expectation of truth: low, medium or high, between subjects) mixed design.

The continuous sampling of the judgment trajectory was also captured. The duration of a statement was split into five equal time points, resulting in a 2 (veracity) x 3 (expectation of truth) x 5 (time point: t_1 to t_5, within subjects) mixed design. Again, the PTJ, A' and B''_D were the dependent measures.

Results

While listening to the statements, participants showed a lie bias when expecting mostly lies and a truth bias when expecting mostly truths. But surprisingly, the effect of base rate beliefs at their final judgment did not mirror the continuous rating of the statement as it was being presented. Raters were most truth biased when expecting mostly truths. Interestingly though, expecting the majority of speakers to lie resulted in approximately half of all statements being judged as truthful.

Base rate information influenced confidence in much the same way. Participants were more confident in their truth than their lie judgments when expecting half or a majority of truths, but were similarly confident in their lie judgments as their truth judgments when expecting mostly lies. Signal detection measures confirm that these effects are attributable to response bias effects rather than accuracy effects.

Belief trajectory. Throughout the duration of each statement raters indicated their belief in the statement. A 2 (veracity, within subjects) x 3 (expectation of truth condition, between subjects) x 5 (time point, within subjects) ANOVA was conducted on the PTJ. In line with the prediction, a main effect of condition was observed, F (2, 57) = 9.64, $p < .001$, $\eta^2 = 0.25$ (Figure 9). Planned post-hoc t-tests found the low truth expectancy condition ($M = .43$, $SD = .12$) differed significantly from the equal split ($M = .53$, $SD = .12$) and high truth expectancy conditions ($M = .60$, $SD = .12$), t (35) = -4.34, $p < .001$, $d = -0.83$, and t (42) = -2.41, $p = .018$, $d = -1.42$, respectively. The equal split and high truth expectancy conditions did not differ significantly, t (40) = 1.69, $p = .096$, $d = 0.58$, although exhibited a medium to large effect size.

The assumption of sphericity was violated for the within-subject effect of time point. A Greenhouse-Geisser correction was used. A main effect of time was found, F (1.99, 113.52) = 7.76, $p = .001$, $\eta^2 = 0.12$, showing significant linear, F (1, 57) = 10.95, $p = .002$, $\eta^2 = 0.16$, and quadratic components, F (1, 57) = 5.51, $p = .022$, $\eta^2 =$

0.09. Finally, there was a main effect of veracity, F (1, 57) = 4.32, p = .042, η^2 = 0.07, such that truths received a significantly higher PTJ (M = .55, SD = .18) than lies (M = .49, SD = .14).

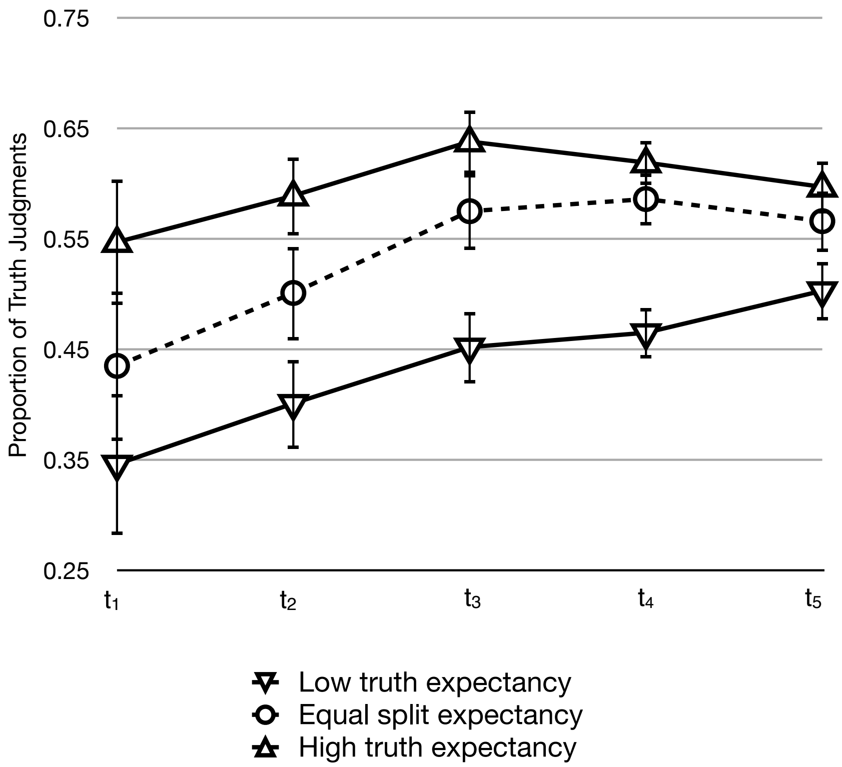

Figure 9. The proportion of truth judgments over the course of the average statement, split by beliefs about the base rate. Whiskers denote standard error.

Signal detection measures were used to separate the effects of accuracy and response bias. Two 3 (truth expectancy condition) x 5 (time point) ANOVAs were conducted on the A' and B"$_D$ scores, respectively. There were no significant effects on A' scores, all ps > .174. The measure of response bias, B"$_D$, found a main effect of time, F (2.28, 129.94) = 7.43, p = .001, η^2 = 0.12, and more importantly for the current purposes, a main effect of condition, F (2, 57) = 12.09, $p < .001$, $\eta^2 = 0.30$. Planned pairwise comparisons reflected those conducted on the PTJ. Specifically, the low truth expectancy condition showed, if anything, a lie bias (M = -.22, SD = .37), which differed significantly from response bias of both the equal split, (M = .16, SD = .37) and high truth expectancy conditions (M = .33, SD = .37), t (35) = -3.09, p = .009, d = -1.49, and t (42) = -1.03, p < .001, d = -1.49, respectively. The equal split and high truth expectancy conditions did not show a significantly different degree of truth bias, t (40) = -1.47, p = .443, d = -0.45.

Finally, three one-sample *t*-tests compared the degree of bias, as measured by B''_D, to no bias, zero. The low truth expectancy condition showed a significant lie bias, t (18) = -2.97, p = .008, d = -0.68. The equal split condition showed no evidence of a bias, t (16) = 1.91, p = .075, d = 0.46. Finally, the high truth expectancy condition showed a significant truth bias, t (23) = 3.88, p = .001, d = 0.79. Bayes factors were calculated for each using JZS one-sample *t*-test equivalent using a Cauchy prior distribution of scaling r = 0.5 over the effect sizes. For the low truth expectancy conditions, the data were 6 times more likely under the hypothesis that there was a lie bias, while for the high truth expectancy the data were 45 times more likely under the hypothesis of a truth bias. However, for the equal split condition, the data were not sufficient to determine whether there was a response bias or a lack of bias, generating a Bayes factor of 1.23. These findings are in line with the hypothesis.

Truth bias as base rate belief. After rating throughout the trial, participants provided a single lie-truth judgment of the statement as a whole. A 2 (veracity, within subject) x 3 (expectation of truth condition, between subjects) mixed ANOVA was conducted on the PTJ. A main effect of veracity was found, F (1, 80) = 64.64, p <

.001, $\eta^2 = 0.45$, such that truthful statements (*M* = .67, *SD* = 0.16) received a higher PTJ than lies (*M* = .50, *SD* = 0.16). The predicted main effect of expectation was also significant, $F(2, 80) = 24.30$, $p < .001$, $\eta^2 = 0.38$. Planned *t*-tests found the high truth expectancy condition received a significantly higher PTJ than the equal split expectancy and low truth expectancy conditions, $t(80) = 3.47$, $p < .001$, $d = 1.32$ and $t(80) = 4.71$, $p < .001$, $d = 1.79$ respectively. The equal split and low truth expectancy conditions, however, did not differ significantly, $t(80) = 1.26$, $p = .079$, $d = 0.49$ (see Table 1).

The Veracity x Expectation interaction was also significant, $F(2, 80) = 3.20$, $p = .046$, $\eta^2 = 0.07$. Bonferroni-corrected post-hoc *t*-tests showed that when rating lies, the high truth expectancy condition produced a significantly higher PTJ than both the equal split and low truth expectancy conditions, both *p*s < .001, *d*s = 1.40 and 1.48 respectively, but the equal split and low truth expectancy conditions did not differ from each other, $t(52) = -0.30$, $p > .999$, $d = -0.09$. Truthful statements received a significantly different PTJ in each of the three conditions, with the high truth expectancy condition receiving a significantly higher PTJ than both the equal split, $t(56) = 2.52$, $p = .031$, $d = 0.69$, and low truth expectancy conditions, $t(55) = 5.26$, $p < .001$, $d = 1.43$, and the equal split condition receiving a significantly higher PTJ than the low truth expectancy condition, $t(52) = 2.61$, $p = .031$, $d = 0.73$. That is, ratings of truths followed the predicted decline in PTJ in line with the raters' expectations. However when rating lies the decline in PTJ from high to low expectancy did not follow predictions, with participants in the low truth expectancy making a comparable degree of truth judgments as those in the equal split expectancy condition.

Table 1. Mean proportion of truth judgments and accuracy rates in each expectancy condition.

	Low truth expectancy		Equal split expectancy		High truth expectancy	
	M	*SD*	*M*	*SD*	*M*	*SD*
Proportion of truth judgments						
*Truths**	.55[a]	0.16	.67[b]	0.18	.79[c]	0.15
*Lies**	.41[a]	0.17	.42[a]	0.18	.66[b]	0.15
Overall	.48[a]	0.13	.55[a]	0.14	.72[b]	0.12
Accuracy						
*Truths**	.55[a]	0.16	.67[b]	0.18	.79[c]	0.15
*Lies**	.59[a]	0.16	.58[a]	0.18	.34[b]	0.16
Overall	.57[a]	0.10	.62[a]	0.10	.56[a]	0.09
Signal detection measures						
A'	.61[a]	0.16	.69[b]	0.13	.62[a]	0.15
B''_D	-0.09[a]	0.49	0.13[a]	0.50	0.70[b]	0.29

Note: Means with a different letter superscript in a row differ significantly ($p \leq .05$) from each other. *Post hoc tests, Bonferroni-adjusted alpha.

Accuracy. The differences in the PTJ when rating lies versus truths may suggest a shift in accuracy rather than a shift in bias. This possibility was explored using a mixed ANOVA on the proportion of correct responses. Accuracy and bias were further delineated using signal detection measures.

A 2 (veracity, within subjects) x 3 (expectancy condition, between subjects) mixed ANOVA was conducted on accuracy rates. All main effects and interactions were significant. The main effect of veracity, $F\ (1, 80) = 31.13$, $p < .001$, $\eta^2 = 0.28$, indicated greater accuracy when judging truths (*M* = .67, *SD* = 0.16) than lies (*M* = .50, *SD* = 0.16). The main effect of condition was also significant, *F* (2, 80) = 3.20, *p* = .046, η^2 = 0.07. The equal split condition was more accurate than the high truth expectancy condition. Whilst a Bonferroni-corrected post-hoc *t*-test found this difference was not significant, *t* (56) = 2.40, *p* = .059, *d* = 0.64, there was a medium to large effect size. The equal split condition was not significantly more accurate than the low truth expectancy condition, *t* (52) = 1.96, *p* = .167, *d* = 0.53, although again showing a medium effect size. The high and low truth expectancy conditions also did not differ significantly, *t* (55) = 0.35, *p* > .999, *d* = 0.11. That the equal split condition was the only group given veridical base rate information may explain these differences.

The Veracity x Expectancy interaction was significant, *F* (2, 80) = 24.30, *p* < .001, η^2 = 0.38. Bonferroni-corrected post-hoc comparisons indicated when judging lies, the high truth expectancy condition was significantly less accurate than both the equal split and low truth expectancy conditions, both *p*s < .001, *d*s = -1.29 and -1.48 respectively, but the low truth and equal split expectancy conditions did not differ significantly, *t* (52) = -0.30, *p* > .999, *d* = 0.08. When rating truths, accuracy was highest in the high truth expectancy condition, which differed significantly from the equal split and low truth expectancy conditions, *t* (56) = 2.52, *p* = .031, *d* = 0.67**,** and *t* (55) = 5.25, *p* < .001, *d* = 1.34, respectively. Also, accuracy was significantly higher in the equal split than in the low truth expectancy condition, *t* (52) = 2.61, *p* = .031, *d* = 0.73. That is, those expecting mostly truths tended to be the most accurate in judging truths, those expecting mostly lies were the most accurate when it came to judging lies, and the equal split expectancy condition, which was given veridical information about the base rates, were generally accurate whether rating lies or truths.

Signal detection measures. To confirm the shift in the PTJ between the expectation conditions was attributable to a shift in bias rather than a shift in accuracy, one-way ANOVAs were performed using A' and B''_D as the dependent variables to explore the effects of accuracy and bias, respectively, independently of each other. There was no evidence that accuracy differed between the three conditions, F (2, 80) = 2.19, p = .119, η^2 = 0.05. As predicted, bias did differ between the three conditions, F (2, 80) = 25.29, p < .001, η^2 = 0.39. The high truth expectancy condition was significantly more truth biased than both the equal split and low truth expectancy conditions, both ps < .001, ds = 1.36 and 1.83, respectively, but the equal split condition was not significantly more truth biased than the low truth expectancy condition, t (52) = -1.89, p = .063, d = 0.52.

Three Bonferroni-corrected *t*-tests compared the A' scores of each of the three conditions to chance rate (0.50). The low, t (25) = 3.39, p < .001, d = 0.69, equal split, t (26) = 7.19, p < .001, d = 1.46, and high truth expectancy conditions, t (29) = 4.24, p < .001, d = 0.80, all showed accuracy significantly above chance rates. Similar analyses conducted on the B''_D scores found both the low truth expectancy, t (25) = -0.95, p = .350, d = -0.18, and equal split conditions, t (26) = 1.39, p = .178, d = 0.26, showed no evidence a significant bias in responding in the final judgment. Only the high truth expectancy condition showed a significant truth bias, t (30) = 13.03, p < .001, d = 2.41.

Bayes factors were calculated to determine whether the response bias showed evidence in favour of no effect or of a present effect. A Cauchy prior distribution with a scaling factor of r = 0.5 was used. An equivalent of a one-sample t-test was performed using the JZS t-test described by Rouder, Speckman, Sun, Morey and

Iverson (2009). When expecting mostly lies, the data were 2.5 times more likely under the null hypothesis, offering moderate support in favour of a lack of response bias in this condition. When expecting an equal split of lies and truths, the data were 1.65 times more likely under the null, suggesting there was not sufficient data to determine whether there was or was not a truth bias in this condition. Finally, when expecting mostly truths there was very strong evidence in favour of a truth bias, with a Bayes factor of approximately 50 billion. Thus there appeared to be a lack of bias when expecting mostly lies but a truth bias when expecting mostly truths at the point of the final judgment.

Divergence in the final judgment. The findings thus far appear to indicate the belief trajectory and the end judgment diverge: the former shows a lie bias when expecting mostly lies, but the latter does not necessarily reflect this, showing no evidence of a bias. The main effect of base rate expectations on the belief trajectory collapses across time, but a main effect of time was observed such that the PTJ was seen to increase. Thus the belief trajectory and the final judgment differences may be accounted for by this change over time.

A linear regression asked whether the belief trajectory at the end of the statement, i.e. at t_5, was able to predict the final judgment after partialling out the effect of base rate beliefs. The purpose of this analysis was to determine whether there truly is a divergence from the end of the belief trajectory to the point of final judgment. The predictors were the PTJ at t_5 and the truth expectancy condition. The outcome variable was the PTJ at the end of the trial. Expectations about the base rate at t_5 significantly predicted the PTJ at the end of the trial, $b = .11$, $t\ (74) = 6.03$, $p < .001$, as did the PTJ at t_5, $b = .34$, $t\ (74) = 3.65$, $p < .001$. That is, the final lie-truth judgment was in part related to the end point of the belief trajectory and on their expectations about the base rate.

Confidence. Replicating much past research (see Aamodt & Custer, 2006), there was no significant correlation found between confidence and accuracy, *r* (81) = 0.06, *p* = .615. It is of interest to ask whether anticipating a majority of truthful statements leads to higher confidence in making a truth judgment, and whether anticipating a majority of lies leads to higher confidence in making lie judgments. A 2 (judgment type: lie or truth judgment, within subjects) x 3 (expectancy condition, between subjects) mixed ANOVA was conducted on the confidence scores when making lie and truth judgments. There was no main effect of condition, *F* (2, 80) = 0.32, *p* = .731, η^2 = 0.01: there was no evidence that participants felt any more or less confident solely as a result of differing expectations. A Bayes factor was calculated using a Cauchy prior distribution over the effect sizes with a scaling factor of r = 0.5. The ANOVA design above served as the complex model and was compared against a model without the main effect of expectancy condition. Both models specified full random effects for participants and speakers. In order to prefer the complex model with a main effect of expectancy condition, prior odds favouring it of at least 2.8 would be needed, offering moderate support for the lack of an effect of base rate expectations.

There was a main effect of judgment type, *F* (1, 80) = 14.48, *p* < .001, η^2 = 0.15. Raters were more confident in making truth judgments (*M* = 4.59, *SD* = 0.81) than lie judgments (*M* = 4.28, *SD* = 0.85), as found in prior research (DePaulo et al., 1997).

In line with the prediction, there was a Judgment Type x Expectancy Condition interaction, *F* (2, 80) = 16.02, *p* < .001, η^2 = 0.29 (Figure 10). Planned *t*-tests provided no evidence that raters in the low truth expectancy condition were more or less confident in making truth than lie judgments, t (25) = -1.85, p = .069, d = -0.32, but both the equal split, t (26) = 2.39, p = .019, d = 0.41, and the high truth expectancy conditions, t (29) = 6.34, p < .001, d = 1.03, were significantly more confident in making truth ratings than lie ratings. Again a Bayes factor was calculated with parameters specified above. The complex model specified the main effect of judgment type using only the data from the low truth expectancy condition, whereas the simple model specified no predictor variables, with both models specifying random effects for participants and speakers. The Bayes factor of 1.1 indicated there was not sufficient power to determine whether there as evidence in favour or against the difference in confidence when making lie or truth judgments in the low truth expectancy condition.

The shift in confidence between conditions was of interest as well as between lie-truth judgments, although here no specific predictions were made. Bonferroni-corrected *t*-tests indicated that, when making lie judgments, the only significant difference was a greater confidence in making lie judgments in the low truth expectancy condition compared to the high truth expectancy condition, t (55) = 1.91, p = .045, d = 0.67, all other *p*s > .179. Similarly, when making truth judgments, there was greater confidence in the high compared to the low truth expectancy condition, t (55) = 2.56, p = .037, d = 0.69, but no other comparisons were significant, *p*s > .213.

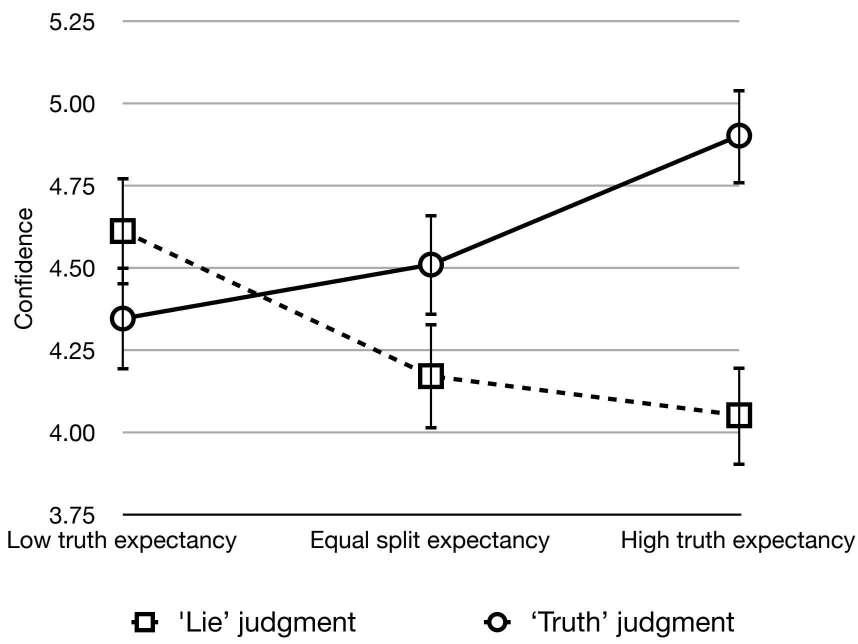

Figure 10. Confidence when making judgments of 'lie' or 'truth' as a function of base rate beliefs. Whiskers denote standard error.

Discussion

People have a somewhat regular propensity to lie, around once or twice per day according to some reports (DePaulo et al., 1996), but the rate of honesty is far greater (Caspi & Gorsky, 2006; Clark & Clark, 1977; DePaulo et al., 1996; Fan et al., 1995; O'Sullivan, 2003; O'Sullivan et al., 1988). Perhaps it is unsurprising, then, that raters tend to judge statements as truthful more often than as deceptive. This study reports on the effects of the perceived base rate of honesty on the tendency to believe others. The findings indicate the truth bias can be moderated by expectations of the base rate of honesty. Across the course of the statement and the raters' developing judgments, there was an effect of base rate beliefs in the predicted direction: raters showed a lie bias when expecting mostly lies and a truth bias when expecting mostly truths. That is, under uncertainty and when there was no other information available to the rater about the current task, they relied on generalised contextual information.

The final judgments of the statements only partially reflected this effect. There was a greater truth bias observed when expecting mostly truths compared to expecting an equal split of lies and truths. Raters in both conditions expressed more confidence in making truth compared to making lie judgments. However, despite expecting every four out of five speakers to deceive them, raters in the low truth expectancy condition surprisingly showed a rate of truth responding that did not differ significantly from the equal split condition. Signal detection measures found no significant response bias when raters believed the majority of statements were to be deceptive. Similarly, they were no more or less confident in their truth ratings than their lie ratings. Expecting a high rate of dishonesty was sufficient to cause a change in confidence compared to the equal split and high truth expectancy conditions, but it was not sufficient to cause a reversal such that raters in the low truth expectancy condition were more confident in their lie than their truth judgments. Put another way, they became appropriately less confident in their truth judgments, but did not have increased confidence in making lie judgments. In each base-rate belief condition, confidence in making truth compared to lie judgments reflected the tendency to exhibit a truth bias in the final judgment.

It might appear the belief trajectory is distinct from the final lie-truth judgment insofar as the degree of truth bias as the statement was being delivered differed from the truth bias observed at the end of the trial. However, this difference can be accounted for by the fact that the degree of truth bias shifted over time. The findings seem to indicate perceptions of the base rate can cause both a lie bias and a truth bias, but that over time there is an increasing tendency towards believing the speaker is being honest.

Beliefs about the base rate may have had a relatively early effect on responding because there was little information available from the speaker at the initial moments. As a result, raters may have utilised other available evidence, in this case their beliefs about the base rate. This initial rating based on base rate beliefs may have served as an anchor (Fan et al., 1995; Zuckerman et al., 1987), from which the belief developed. Over time more behavioural information would have become available from the speaker. Given that both the honest and deceptive speakers seek to be perceived as telling the truth (DePaulo et al., 2003), their self-presentational behaviour may have successfully persuaded raters of their honesty and resulted in an increase in the PTJ. That is, the shift towards a truth-biased response may be attributable to increasingly availability of apparently honest behaviour from the speaker (Buller & Burgoon, 1996; Chung & Fink, 2008).

An alternative account would suggest the initial use of base rate information is followed by the onset of some cognitive process that causes the truth bias. This account suggests the shift towards a truth bias is a cognitive phenomenon on part of the rater rather than attributable to the self-presentational behaviour of the speaker, and that it operates at a later time scale than base rate information. Whether a behavioural or cognitive account can best explain the shift in the PTJ over time is unclear.

The findings could be explained with sole reference to the use of base rate information, provided a distinction is made between the sample base rate and the population base rate. By sample base rate we mean the information offered to participants in this study regarding the small sample of speakers they were to rate. This varied between conditions. By population base rate it is meant the information participants bring to the task about the base rate of honesty from their entire history of interactions in their daily lives. As discussed, the rate of honesty in real-world interactions is far greater than the rate of deception experienced (e.g. Caspi & Gorsky, 2006; DePaulo et al., 1996; O'Sullivan et al., 1988). As a statement progresses, raters may be more influenced by their experiential knowledge of the population base rate and shift towards a truth belief. Future research should seek to contrast such cognitive accounts with behavioural accounts of the truth bias.

The findings suggest a sceptic view may result in a tendency towards disbelieving whilst listening to a statement, in contrast to Gilbert et al. (1990, Study 2; Gilbert, 1991) but in line with other research (Masip, Alonso, et al., 2009; Mayo et al., 2004; Nieuwland & Martin, 2012). Yet at the point of making the final judgment there is no evidence of a lie bias. One possible interpretation is that a bias brought to the task can be amplified or dampened by perceptions of how often people in general will tell the truth. Base rate information is incorporated into the judgment and can modulate the degree of the response bias. An important question for future research, then, is whether police investigators who bring a lie bias to the task will show an increased lie bias if they believe most of the speakers will lie, and a dampened response bias if they believe most speakers will tell the truth.

This study offers the first empirical exploration of the effects of perceived base rates on the truth bias. Interestingly, whilst over the course of the statement their developing judgment was in line with the perceived base rates, by the point of their final judgment there was no direct one-to-one relationship between the expected proportion of deceptive and honest statements and the actual proportion of lie and truth judgments. An increasing shift towards believing the statement to be true over time was observed prior to the final judgment, although it is unclear why this shift occurred. This experiment shows beliefs about the base rate (independent of the actual base rate) do influence the trajectory of the belief and the final lie-truth judgment, resulting in an early lie bias when expecting mostly lies and an early truth bias when expecting mostly truths. Taking into account the perceived base rate is better seen as a normative use of information (Nisbett & Ross, 1980) rather than a biased and inaccurate view of the world.

One last note of consideration is worthy of mention. The manipulation explicitly informed participants the percentage of truth-tellers in the stimulus set. One might wonder whether participants given this information were acting in line with what they thought was the experimental manipulation, which was clearly presented to participants. That is, there is a worry the results may be an artefact of demand characteristics.

This is of course a valid concern that cannot be dismissed with the current experimental setup. However, it is worth noting that there is no clear evidence of a direct one-to-one mapping between the base rate manipulation and the proportion of truth responses made. Although there was clear early evidence of an effect of base rate information, it is clear from Figure 9 that the PTJ increased over time. In addition, at the point of the final judgment raters showed a propensity towards believing over disbelieving the speaker, albeit to different degrees, regardless of the base rate information given. Of course, despite this complicated pattern of responding, it is not possible to rule out demand characteristics interacting with other factors, and this should be borne in mind when interpreting the current results.

Experiment 7: A Cartesian Mind – Post Comprehension

The findings of Experiment 5 show there is no automatic tendency to believe during the process of comprehension. Rather, the truth bias is invoked by forcing a decision from a rater who has yet to receive enough information from the environment to make a decision. That is, the bias results from the structure of the task rather than from the structure of the mind. Experiment 8 capitalises on this. Here it is considered whether the truth bias observed in the final decision, rather than across the course of the decision making process, can also be modulated by whether the rater is forced into making a judgment.

There is room for much uncertainty even after all the information has been considered. We are particularly good liars and produce few cues to our deception (Bond & DePaulo, 2008; DePaulo et al., 2003; Sporer & Schwandt, 2006, 2007; Vrij, Granhag & Mann, 2010), possibly none at all (Levine, 2010; although see Duran, Dale, Kello, Street & Richardson, 2013, for evidence of nonverbal indicators of deception at the individual level). How a rater should interpret a single instance of a cue is unclear: even truth tellers can appear nervous, for example, resulting from the accusatory situation (Ekman, 1992; Vrij, 2008; Vrij, Granhag & Porter, 2010). There is a degree of uncertainty inherent in the lie detection task, reflected in consistently low accuracy rates (Aamodt & Custer, 2006; Bond & DePaulo, 2006; Kraut, 1980; see also Aamodt & Custer, 2006; Frank & Feeley, 2003).

It is predicted that when able to indicate uncertainty, there will be no need to rely on prior knowledge to make an informed guess and as a result there will be a reduced truth bias.

Method

Participants. Eighty-two University College London psychology students participated in this study. One participant retrospectively withdrew consent. Two participants failed to provide their age and sex, but were not excluded from analyses. Of the 79 participants that reported demographic details, 54 were female (age M = 19.87, SD = 1.31, range 18 to 22). Participants received either course credit or £3 compensation.

Materials, Design and Procedure. The Bloomsbury Deception Set (Street et al., 2011, April) in Experiment 5 was used. Videos were presented and responses collected using MATLAB software (The Mathworks inc., 2000: Natick, MA) and the Psychophysics Toolbox (Brainard, 1997).

The independent variables were the response conditions (LT or LTU) and the veracity of the speakers' statements. The dependent variable was the proportion of truth judgments, resulting in a 2 (response condition: LT or LTU, between subjects) x 2 (speaker veracity: lie or truth, within subjects) mixed design.

The procedure follows Experiment 5 with two exceptions. First, raters did not indicate their belief as the video progressed. Rather, they viewed the video and only afterwards made a response. Second, only raters in the lie-truth (LT) condition (n = 39) made a forced binary choice: raters in the lie-truth-unsure (LTU) condition (n = 40) were given the additional option of indicating their uncertainty. That is, raters viewed a video passively, and after each video either made a lie-truth or lie-truth-unsure judgment.

Results

Having the option of explicitly indicating uncertainty did affect the truth bias, but in the opposite direction than predicted. Being able to indicate uncertainty *increased* the truth bias. Further exploration of the data indicated this effect was due to a shift away from making lie responses and towards making unsure responses. This was further supported by the fact that lies were more likely to receive unsure ratings than were truths.

As in Experiment 5, the unsure responses were removed from the LTU condition in order to compare the LT and LTU conditions. After removal, a PTJ of 0.50 would indicate no bias. A 2 (response condition: LT or LTU) x 2 (veracity: lie or truth) mixed ANOVA with repeated measures on the second factor was conducted with the PTJ serving as the dependent variable. A main effect of veracity was found, F (1, 77) = 36.02, $p < .001$, $\eta^2 = .319$. Truths ($M = .66$, $SD = .18$) received a higher PTJ than lies ($M = .50$, $SD = .19$), indicating raters were above chance accuracy overall. Central to the hypothesis, there was a main effect of response condition, F (1, 77) = 4.50, $p = .037$, $\eta^2 = .055$. Surprisingly, there was a greater bias in the unforced choice LTU condition ($M = .61$, $SD = .15$) than in the forced choice LT condition ($M = .54$, $SD = .14$).

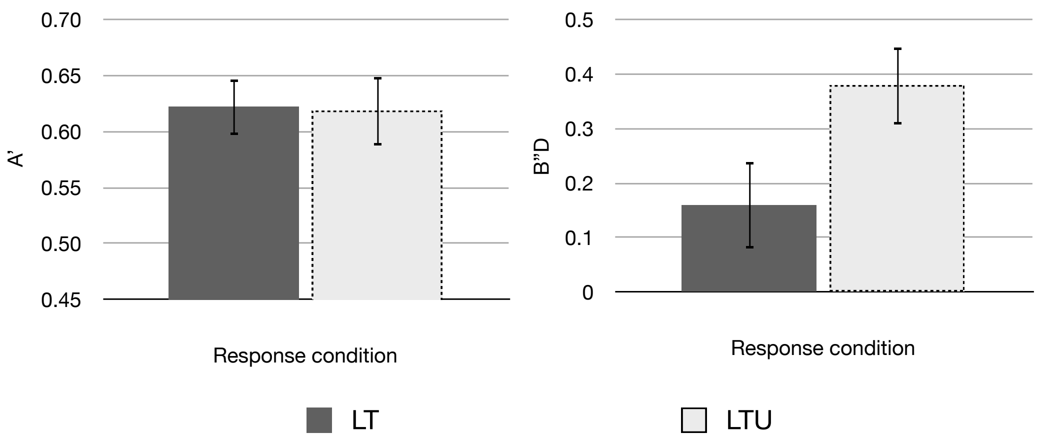

Figure 11. (a) Accuracy (A') in the forced choice LT and the unforced choice LTU conditions. 0.5 indicates chance accuracy (left). (b) Response bias (B''_D) in the LT and LTU conditions. Zero indicates no bias; positive values indicate a truth bias. Whiskers denote standard error.

Signal detection measures of accuracy (A') and response bias (B''_D) were also calculated to determine their independent effects. Two independent samples t-tests found no significant different in accuracy between the LT and LTU conditions, t (77) = 0.09, p = .929, d = 0.02 (Figure 11a), whereas the difference in bias was statistically significant, t (77) = 2.06, p = .043, d = -0.65. (Figure 11b).

Exploration of the data. In order to understand the reversal of the effect, further examination was carried out. First, the unsure responses were reintroduced into the LTU condition data. Comparing the proportion of truth judgments between the conditions and the proportion of lie judgments between the two conditions will give an indication of whether the unsure responses in the LTU condition were likely to draw away from making lie responses. I will refer to this as the *artefactual truth bias* because the proportion of truth to lie judgments increases only as a result of making fewer lie responses. Another possibility is that the unsure response is not used frequently, and the proportion of truth responses in the LTU condition exceeds the proportion of truth responses in the LT condition even when the unsure responses are considered. I will refer to this as the *veridical truth bias* to indicate the fact that truth responding is greater in the LTU condition than in the LT condition when all responses are taken into account.

Two independent samples *t*-tests compared the proportion of truth judgments and the proportion of lie judgments in the two response conditions. The proportion of truth responses did not differ significantly between the LT (M = .54, SD = .14) and the LTU conditions (M = .50, SD = .14), t (77) = 1.23, p = .223, d = 0.28. The proportion of lie responses did differ significantly, t (77) = 4.25, p < .001, d = 0.96. There were significantly fewer lie judgments made in the LTU condition (M = .33, SD = .14) than in the LT condition (M = .46, SD = .14). This suggests an artefactual truth bias: it was not the case that raters were more likely to make truth judgments in the LTU condition, but rather that they were less likely to make lie judgments, instead making use of the unsure response (Figure 12).

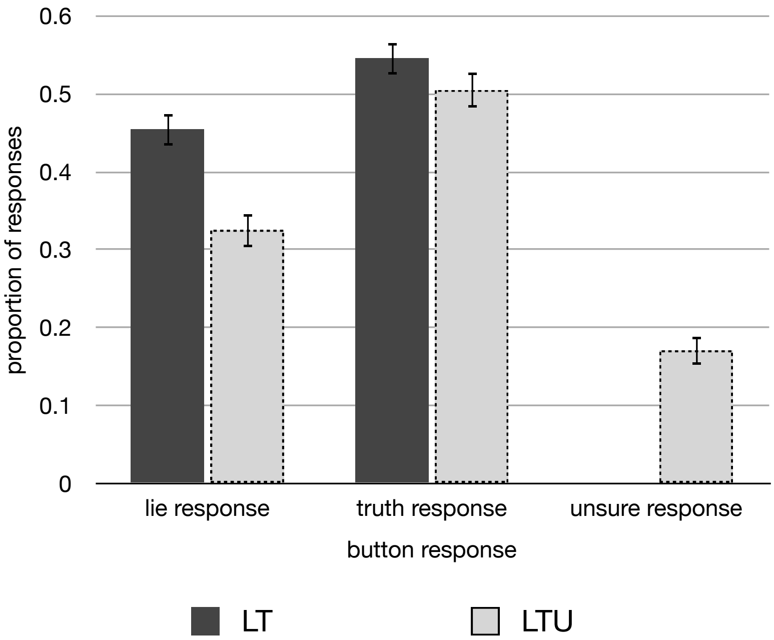

Figure 12. The proportion of lie, truth, and unsure (where applicable) responses in the LT and LTU conditions. As can be seen, the increased truth bias in the LTU unforced choice condition results from a reduction in the use of the 'lie' response. Whiskers denote standard error.

A second analysis explored the use of the unsure response in the LTU condition in more detail. Prior research suggests there is greater uncertainty when judging lies than judging truths (e.g., DePaulo et al., 1997). A paired-samples *t*-test compared the proportion of unsure responses in the LTU condition when rating lies versus when rating truths. As expected, there was a significantly greater proportion of unsure responses when rating lies (M = .39, SD = .19) than when rating truths (M = .26, SD = .18), t (38) = 3.40, p = .002, d = 0.15.

Discussion

The current experiment sought to test the possibility whether the truth bias observed in the final judgment could be reduced or removed by allowing respondents to indicate their uncertainty explicitly, rather than forcing them into a binary choice. Contrary to expectations, there was a greater truth bias when given the option of indicating uncertainty. Further exploration of the data revealed the increase in the truth bias was artefactual: there was no increase in truth responding in the LTU condition, but rather there was a decrease in making lie responses and instead favouring the use of the unsure response.

It is important to note that a Cartesian system does not make predictions regarding the truth bias after comprehension, only that there is no decision made during the comprehension process (Gilbert, 1991). The findings of this study cannot be said to contradict the Cartesian account. However, they may be explained as an adaptive decision-maker that makes use of contextual knowledge.

Of particular interest is why the truth bias increased in the unforced LTU condition. Raters were less likely to use the lie response and instead opt for the unsure response. This suggests that, in the case of a forced binary choice, when there is uncertainty this in itself can be used as a guide towards a judgment: if unsure, guess ‘lie’. This strategy would in general be successful: research indicates people have greater uncertainty when rating a deceptive statement than an honest statement (Anderson et al., 2002; DePaulo, 1992; DePaulo et al., 1997; DePaulo et al., 2003; Anderson, 1999, cited by DePaulo & Morris, 2004; Hartwig & Bond, 2011; Levine et al., 1999; see also Hurd & Noller, 1988). This may be because liars, through fear of getting caught, may purposely deliver vague statements (Vrij, 2008), or because raters are less familiar with lies than truths (DePaulo, 1992; DePaulo et al., 1996). The current study confirms this view: raters were indeed more likely to make an unsure response when listening to lies than when listening to truths. Thus the apparent increase in bias may be better thought of as a reflection of an adaptive decision maker that can make use of contextual knowledge (lies are typically harder to detect than truths: Levine et al., 1999) by using their internal uncertainty to improve the quality of the judgment (see Gigerenzer & Selten, 2001; Payne et al., 1993).

Adaptive use of uncertainty could only be successful after the information has been comprehended. Being unsure in this respect means the evidence gathered from the environment is not sufficiently clear to make a distinction between whether the statement is a lie or a truth. Because truths are easier to detect than are lies, a phenomenon known as the veracity effect (Levine et al., 1999), internal uncertainty can be taken as an indication of the possibility of the statement being a lie: deceptions are more difficult to detect.

The internal uncertainty claim assumes raters are meta-cognitively aware of their difficulty in making a categorical judgment with the information available, and can use that uncertainty itself as a guide towards making a judgment. This may seem to suggest that with greater uncertainty we may expect raters to be better able to make a decision, provided they can employ their own uncertainty as a factor in the decision process. This claim is not being made; there is a step missing. It is *because the experimental design requires participants make a decision* that they may come to use that uncertainty when forced into judgment. When this necessity to make a judgment is removed, as seen in the LTU condition of this experiment, raters are less likely to make use of their uncertainty in this way. Instead, they prefer to explicitly indicate their uncertainty. That is, indecision itself is used as a factor in the decision when there are additional requirements made by the task, viz. to make a categorical decision.

Uncertainty during comprehension, on the other hand, is due to the lack of information obtained from the environment. Being unsure in this respect results from a lack of information. This *external* uncertainty cannot be used as an indication of deceit or honesty: it is equiprobable that the statement is a lie as it is a truth. The rater has to rely on information that does not relate to the immediate environment, but rather on more general principles from past experience with similar situations, such as the availability heuristic ('people usually tell the truth': Experiment 6).

One weakness with the internal-external uncertainty explanation is that it results from an exploratory analysis of the data. Experiment 8 takes a confirmatory approach to the question, beginning from the hypothesis that judgments made after information has been gathered will show a greater truth bias if not forced into a judgment, an artefact of not making use of internal uncertainty as a guide to deception and instead explicitly indicating uncertainty, thereby reduced lie responses. When task-specific information in the environment has not been encoded, either because sufficient information or time is not available, prior experience with internal uncertainty cannot be used as a guide to deception. Instead, knowledge built up through past experience with similar situations can be used, such as 'people usually tell the truth'.

Experiment 8: The Cartesian Hopi Word Experiment

The aim of the current experiment is to replicate and extend the findings and conclusions of Experiments 5 and 7. After information from the environment has been encoded, it is claimed the use of internal uncertainty can be used strategically as an indication of falsity. If there is no information available in the environment, (external) uncertainty cannot be used strategically. Instead the rater must rely on other available information, such as that built up from past experience with similar situations.

The current study closely replicates an experiment carried out to support the Spinozan view (Gilbert et al., 1990, Study 1) for two reasons. First, it is important to show that previous data supporting the Spinozan view can better be accounted for by an informed Cartesian view that takes an adaptive approach to decision-making. Second, the experimental setup lends itself well to the empirical testing of the internal-external account suggested by the findings of Experiment 7.

Gilbert et al. (1990, Study 1) manipulated the amount of time for encoding veracity information (for a fuller description of their study, see the methods of this experiment). Participants were set the task of learning the definitions of ostensibly Hopi Indian words, but were in fact nonsense words (e.g. 'A twyrin is a doctor'). After presentation of the word-definition pair, a true or false signal word was presented to indicate whether the definition was either true (a twyrin is in fact a doctor) or false (a twyrin is not a doctor). The veracity signal word was presented for either 3 sec or interrupted after only 750 msec. When given little time for encoding, raters showed a truth bias, which the authors took as confirmatory support for the Spinozan view. However, as was argued in Experiment 5, the evidence is compromised by the fact that raters were forced to decide under external uncertainty: whether due to a lack of available information or, in this instance, a lack of time to process the available information. Under external uncertainty, raters can make use of their prior knowledge from similar situations, such as 'people usually tell the truth' (Grice, 1975; O'Sullivan et al., 1988), and an understanding that communication of new information needs to be *true* information if it is to be useful (see Fiedler et al., 1996; Grice, 1975; Swann et al., 1982). This conclusion was supported in Experiment 5.

Given sufficient time to process the information, i.e. without interruption, the researchers found a reduced truth bias (Gilbert et al., 1990, Study 1). The findings of Experiment 7 suggest that, given task-specific information has been gathered, uncertainty reflects indecision about the information (internal uncertainty). Rather than guessing randomly, internal uncertainty can be used adaptively. It has been known for a long time that processing negated statements takes longer and they are more difficult to comprehend than non-negated statements (Anderson et al., 2010; Carpenter & Just, 1975; Clark & Chase, 1972; Glenberg et al., 1999; Johnson-Laird, 1983; Trabasso et al., 1971). What is more, true statements are encountered more often than negated statements in our daily lives (Skurnik, 1998, cited by Hasson et al., 2005). Thus uncertainty can be taken as an adaptive heuristic guide to the veracity of the statement by assuming those statements that are difficult to process are more likely to be false. This adaptive use of internal uncertainty would be expected to offset the tendency towards accepting statements as true, resulting in a reduced response bias.

This setup allows for a replication of the Spinozan effect as shown by Gilbert and colleagues (1990) and to further demonstrate how the adaptive use of either past experience (when little task-specific information is available) or internal uncertainty (when task-specific information fails to guide the decision to a conclusion) can result in biased or unbiased responding.

Participants were led to believe they were taking part in a learning experiment and that their task was to learn words from the Hopi Indian language. Definitions were presented and followed by signal word indicating whether the definition was true or false. The amount of time processing the signal word was varied, replicating Gilbert et al. (1990, Study 1). After the learning phase followed a test phase, and half the participants were asked to recall whether the presented definition was true or false for the word presented, or whether the word-definition pair had not been seen during learning. The other half of the participants had the additional option of explicitly indicating they were unsure about the statement's veracity.

It was hypothesised that those forced to make a true-false judgment would exhibit the Spinozan truth bias, replicating Gilbert et al. (1990, Study 1). That is, a truth bias should be exhibited when processing the veracity of the statement during the learning phase was interrupted; there would be a reduced biased if there was no interruption during learning. When able to indicate uncertainty, the adaptive decision-maker account predicts a *reverse* Spinozan effect. It is claimed that interruption in this task would lead to the use of prior knowledge, but only when forced. If unforced, biased responding should reduce because the decision-maker is no longer forced to make a judgment under external uncertainty, in line with the reasoning of Experiment 5. Given sufficient time to encode the veracity information, uncertainty can be used adaptively. The processing of false statements is considered more difficult than processing affirmed statements (Carpenter & Just, 1975; Clark & Chase, 1972; Johnson-Laird, 1983; Trabasso et al., 1971), perhaps because they are underspecified: a negated statement ('a twyrin is not a doctor') gives no information as to the true meaning of the word, and as such may require the processing of the various alternatives (Anderson et al., 2010; Glenberg et al., 1999). When able to indicate uncertainty, there would be no necessity to make an informed guess if unsure. With fewer 'false' guesses, an artefactual truth bias is predicted.

Method

Participants. Eighty-three participants took part in return for either course credit or £3 compensation. One participant took a phone call mid-experiment. The delay manipulation explained below did not work correctly for three participants. These four participants were excluded. Of the remaining 79, 53 were female (age M = 21.47, SD = 4.56, range = 18 to 41). Participants were pre-screened for their knowledge of the Hopi Indian language. All participants spoke English as their native language.

Materials. Twenty-eight stimulus propositions were taken from Gilbert et al. (1990, Study 1, Table 1). Two additional word-definition pairings were generated in order to balance the design, another four were generated for the practice phase and a final for use during a practice phase, and a further nine foil items were generated for use during the test phase. Table 2 lists the stimulus items presented during both the learning and test phases.

Procedure. *Learning phase.* Participants signed up to take part in a 'natural language learning' experiment. Instructions were adapted as closely as possible from Gilbert et al. (1990, Study 1). A series of propositions were presented one at a time of the form 'An X is a Y' for a period of 8 s, where X was a nonsense word and Y a noun. The order of presentation was fixed as shown in Table 2. A blank screen of either 2 or 10 s followed each presentation, as per Gilbert et al. (1990, Study 1). It was explained this gap was because 'the computer is attempting to locate a new proposition in its internal dictionary for the next trial'.

On two-thirds of the trials, a signal word was then presented, indicating whether the definition was accurate of the word (true) or not (false) for 3 s. It was explained that in order to simulate natural learning a signal word would not always be given because ‘in natural language learning we do not always get feedback about the meaning of words.’ On the remaining third of the trials, there was no signal word and instead a second blank screen was shown for 3 s. The counterbalancing presented in Table 2 meant that each proposition was followed by either a true, false or no signal word across the course of the experiment.

A second simultaneous task was given to participants. As well as the word-learning task, participants had to response to a 600 Hz tone as fast as possible with a button response. Instructions explained ‘some research suggests that the speed of response is related to the ability to learn new languages’; however the true purpose of the tone response was to interrupt processing of the signal word. Whether participants actually responded to this tone was captured, but because the purpose of the tone was simply to interrupt participants, the responses are not analysed here. Of the 30 propositions, the first and last six were used as buffers to avoid primacy and recency effects. Of the remaining 18 trials, 12 were followed by a signal word of true or false. Interruption occurred on the six emphasised propositions in Table 2: Gilbert and colleagues (1990) interrupted only four of the propositions. The reason for this change was to balance the number of interrupted and uninterrupted propositions that would be used in the test phase.

Table 2. Three orders of stimulus item presentation during the learning phase of the Uncertain Hopi Experiment.

Trial	Proposition	Order			Trial	Proposition	Order		
		1	2	3			1	2	3
1	A volike is a puddle	T	B	F	***16***	***A suffa is a cloud***	***T***	***B***	***F***
2	A ghoren is a jug	T	B	F	17	A walive is a bear	B	F	T
3	A monishna is a star	F	T	B	18	A tecrill is a mouse	F	T	B
4	A cirell is a tree	B	F	T	19	A basol is a fisherman	F	T	B
5	A tarka is a wolf	T	F	B	20	A casin is a rope	B	F	T
6	A dinca is a flame	B	T	F	21	A nasli is a snake	T	B	F
7	***A polay is a stream***	***B***	***F***	***T***	***22***	***A twyrin is a doctor***	***T***	***F***	***B***
8	A tica is a fox	T	B	F	23	A bandi is a raccoon	B	T	F
9	A bilicar is a spear	F	T	B	***24***	***A dalith is a root***	***F***	***B***	***T***
10	A korrom is a mountain	F	T	B	25	A tiloom is a cup	F	B	T
11	***A curira is a necklace***	***B***	***T***	***F***	26	A gafin is a pinecone	F	T	B
12	A waihas is a fish	F	B	T	27	A hib is a canoe	B	F	T
13	A rotan is a hunter	T	B	F	28	A trica is a weasel	T	B	F
14	***A wika is a deer***	***F***	***T***	***B***	29	A neseti is a bee	T	F	B
15	A rirg is a valley	F	B	T	30	An eprata is a berry	B	F	T

Note. T = true, F = false, B = blank screen. Four of the ***bold italicised*** items were interrupted on one run of the experiment.

Test phase. The first and last six items from the learning phase were removed to counter potential primacy and recency memory effects. This left 18 trials, of which 12 were followed by either a true or false signal word. Half of these 12 word-definition pairs were followed by a 2 s gap before the presentation of the signal word, whilst the remaining pairs were followed by a 10 s gap. A further nine foil items that were not presented during learning were added to the test phase, resulting in a total of 21 test items. Item order was randomised.

A written instruction was presented before each test phase, at which point the experimenter re-entered the room to paraphrase the instruction. Those in the false-true (FT, n = 43) response condition were given three response options: they could either indicate the proposition was false, true, or that the proposition had not been presented during learning. Participants in the false-true-unsure (FTU, n = 36) condition were given these three responses in addition to a fourth ‘unsure’ response. Importantly, it was stressed to participants in all conditions that they should only use the ‘not seen’ response if they believe it was a new proposition: if they recalled seeing the proposition but could not recall whether it was true or false, they should not use the ‘not seen’ response and to use one of the remaining two or three response options available (depending on whether the participant was in the FT or FTU experimental condition).

Results and Discussion

Proportion of true judgments. The nine foil items in the test phase were removed from analysis, leaving 12 critical stimuli for analysis. Of these, half were interrupted during the presentation of the signal word in the learning phase, meaning they had only a short space of time in which to encode the relevant veracity information. Not-seen and unsure responses were removed from analyses.

A 2 (veracity: true or false, within subjects) x 2 (interruption: interrupted or uninterrupted during learning, within subjects) x 2 (response condition: FT or FTU, between subjects) mixed ANOVA was conducted on the PTJ. A main effect of veracity was found, F (1, 73) = 4.87, p = .030, η^2 = 0.06[6], with a greater PTJ for true (M = .70, SD = .29) than false (M = .58, SD = .32) statements. No other main effects were significant, all ps > .380.

A Veracity x Interruption interaction is predicted by the Spinozan mind hypothesis: false items should be mistaken as true only if interrupted during learning. A significant interaction was found, F (1, 73) = 14.48, p < .001, η^2 = 0.17. When interrupted during learning, participants were equally likely to judge false statements as true (M = .65, SD = .38) as they were to (correctly) judge true statements as true (M = .65, SD = .40), t (74) = -0.02, p = .979, d < .001. When uninterrupted during learning, raters were more likely to judge true statements as being true (M = .75, SD = .27) than they were false statements (M = .51, SD = .35), t (74) = -4.45, p < .001, d = 0.77. This provides initial support for a Spinozan view, although it is unclear whether raters were becoming more accurate when they were not interrupted during the learning phase.

More pressing for the current research is whether interruption affected the degree of truth responding in the FT and FTU response conditions. An Interruption x Response Condition interaction was observed, F (1, 73) = 8.14, p = .006, η^2 = 0.10, as shown in Figure 13. The means suggest a crossover effect as predicted by the adaptive decision-maker account. Bonferroni-corrected t-tests found interruption during encoding resulted in significantly more truth judgments than when uninterrupted in

[6] Some participants had missing cells for some of the analyses presented in this section. For example, some participants either indicated they were unsure or that they had not seen the items, which would result in blank cells when calculating the proportion of true responses out of all the false and true responses made.

the FT condition, t (43) = 2.66, p = .010, d = 0.42, as predicted by the Spinozan mind account. In the FTU condition, however, whilst the means suggest a reversal of the effect, the difference in the PTJ between interrupted and uninterrupted items was not significant, t (31) = -1.46, p = .144, d = -0.27. In addition, there was no statistically significant difference between the FT and FTU conditions on those items interrupted during learning, t (74) = 0.77, p = .446, d = 0.18, but the FTU condition showed a greater truth bias on uninterrupted items, t (74) = -2.27, p = .025, d = -0.54. Whether interrupted or uninterrupted during learning, FTU raters exhibited a marked truth bias of around 60% to 70%, whilst the FT raters exhibited the expected Spinozan mind effect.

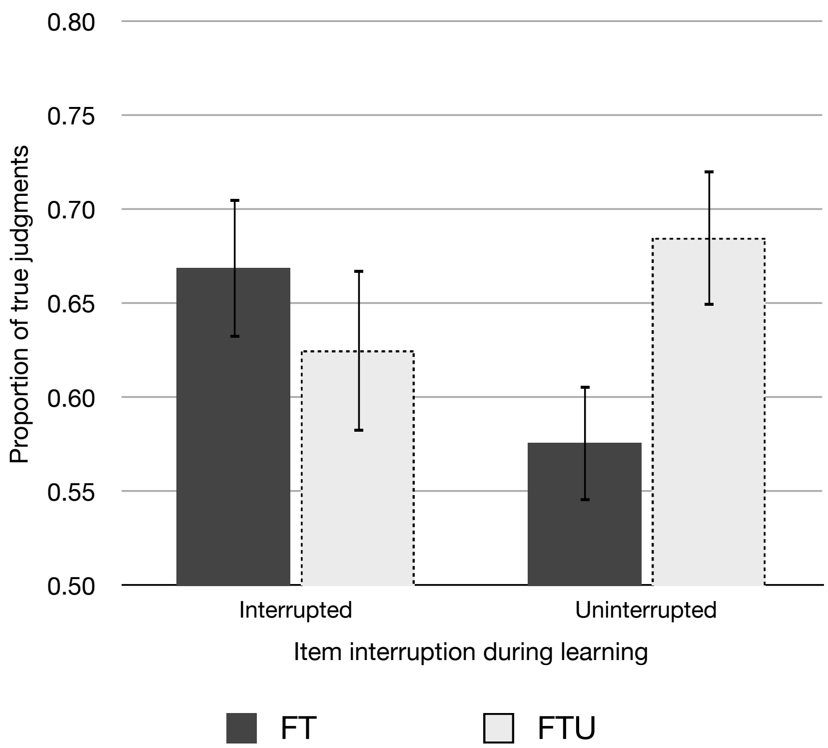

Figure 13. The effect of interruption on the proportion of truth judgments as a function of the response options available: either a forced choice (FT) or unforced choice (FTU). Whiskers denote standard error.

Signal detection measures. The above analyses are presented using the raw proportion of true judgments. These analyses allow for a more ready comparison with those of Gilbert and colleagues (1990, Study 1). However, response bias and accuracy are confounded in these raw measures. Here I will consider the effect of the manipulations on accuracy (A') and bias (B''_D) independently.

A 2 (interruption) x 2 (response condition) ANOVA was conducted on the A' accuracy scores. There was a main effect of interruption, $F(1, 73) = 11.76$, $p = .001$, $\eta^2 = 0.14$: raters were significantly more likely to make an accurate response when uninterrupted during learning ($M = .64$, $SD = .28$) than when interrupted during learning ($M = .51$, $SD = .36$). Thus the act of interruption during encoding was liable to cause inaccuracies.

Conducting the same ANOVA on the B''_D response bias scores reflects the Interruption x Response Condition interaction found using the PTJ scores, $F(1, 73) = 5.79$, $p = .019$, $\eta^2 = 0.07$. Bonferroni-corrected *t*-tests reflect the Spinozan effect in the FT condition and the sustained truth bias in the FTU condition as exhibited in the ANOVA conducted on the PTJ and as illustrated in Figure 13. A series of four one-sample *t*-tests on the B''_D scores determined the FT condition showed a significant truth bias on interrupted items ($M = .46$, $SD = .69$), $t(42) = 4.36$, $p < .001$, $d = 0.67$, and on uninterrupted items, albeit smaller ($M = .22$, $SD = .61$), $t(42) = 2.38$, $p = .022$, $d = 0.36$. The unforced FTU condition showed a truth bias on both interrupted ($M = .34$, $SD = .69$), $t(31) = 2.79$, $p = .009$, $d = 0.49$, and uninterrupted items ($M = .48$, $SD = .64$), $t(31) = 4.47$, $p < .001$, $d = 0.75$.

Thus the FT condition exhibited the Spinozan effect of a truth bias on interrupted items and a reduced (albeit present) truth bias on uninterrupted items. In contrast, the FTU unforced condition did not show the Spinozan effect. The means suggested a reversal of the effect, as predicted, however this failed to achieve statistical significance. Both interrupted and uninterrupted items showed a truth bias, with the two showing no statistically significant difference.

Further exploration. To attempt to understand the sustained truth bias in the FTU unforced choice condition, further analyses were conducted.

Forgetting effect. As alluded to above, interruption resulted in lower accuracy rates. It may be that interruption caused people to forget. Gilbert et al. (1990, Study 1) rejected the forgetting-effect explanation and showed no difference in the tendency to misremember interrupted (9%) versus uninterrupted (8%) items as new items that had not been presented during learning. A 2 (veracity) x 2 (interruption) x 2 (FT/FTU response condition) ANOVA was conducted on the proportion of not seen responses. This proportion was calculated as the sum of the not seen responses divided by the sum of the true, false and not seen responses. A main effect of interruption was found, F (1, 77) = 11.77, p = .001, η^2 = 0.13. Participants were significantly more likely to indicate they had not seen an item if encoding of the veracity information had been interrupted (M = .15, SD = .21) than if there was no interruption during learning (M = .07, SD = .11). There was a non-significant effect of veracity, F (1, 77) = 3.41, p = .069, η^2 = 0.04. Thus the act of interrupting led to forgetting, in contrast to the findings of Gilbert et al. (1990, Study 1). However, those authors did not explore how interruption interacted with veracity of the proposition.

The interruption main effect was qualified by a three-way interaction between veracity, interruption, and response condition, F (1, 77) = 6.64, p = .012, η^2 = 0.08. Post-hoc Bonferroni-corrected *t*-tests were conducted. In the FT condition, interrupted false propositions (M = .22, SD = .27) were significantly more likely to be considered as new items compared to interrupted true propositions (*M* = .09, *SD* = .24), t (42) = 3.57, p = .001, d = 0.51. For uninterrupted propositions, there was no significant difference in the proportion of not seen responses between true (*M* = .06, *SD* = .13) and false items (*M* = .09, *SD* = .15), t (42) = 0.17, p = .869, d = 0.21. Thus in the FT condition, where the replication of the Spinozan truth-bias effect was anticipated, it appears as though there was greater forgetting for false items when interrupted than any other items, shown in Figure 14a.

Gilbert et al. (1990, Study 1) present the proportion of true and false responses for each of the true and false statements in their Figure 2. Although the exact mean values are not reported, the figure indicates approximately 35% of the false statements were accurately judged as false when interrupted, and approximately 35% of the false statements were incorrectly classified as true when interrupted. This leaves approximately 30% of the responses unaccounted for. Given that the only other response available to raters was 'not seen', it seems reasonable to suggest the missing values in the figure represent the not-seen responses. However, caution must be taken in making this assumption. Nonetheless, approximations of the missing values from the Gilbert et al. (1990, Study 1) figure are plotted in Figure 14b.

As can be seen in Figure 14, there is a resemblance between the not seen responses in the current experiment and the unaccounted-for responses in Gilbert et al. (1990, Study 1). Of course whether the differences shown in Figure 14b are statistically significant is not known. What can be said from the findings of the current study is that participants were more likely to judge interrupted false (versus true) propositions as having not been seen during learning. Thus the Spinozan truth-bias effect may be an artefact of increased forgetting of interrupted false statements that leads to a guessing bias, as Gilbert and colleagues (1990) had considered, but rejected. Without an explicit analysis of the raw data from their research, it is unclear whether the Spinozan effect found in their research can also be accounted for as a greater forgetting of interrupted *false* propositions, but it would be worth exploring. The researchers no longer have the data available to explore this possibility (D. Gilbert, personal communication, 30 August, 2013; D. Krull, personal communication, 30 August, 2013; P. S. Malone, personal communication, 30 August, 2013).

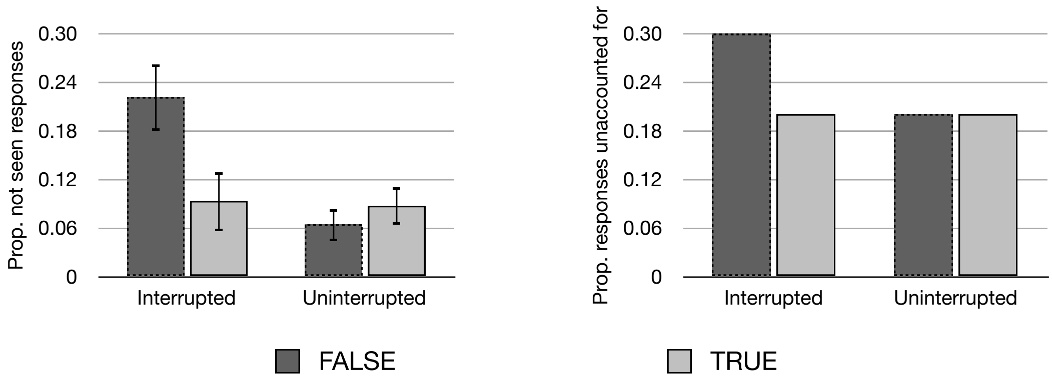

Figure 14. (a) The proportion of not-seen responses in the FT condition of the current study. Raters were more likely to forget interrupted false propositions than any other proposition type, suggesting the Spinozan effect reflects forgetting of false statements when interrupted (left). Whiskers denote standard error. (b) The approximate proportion of responses unaccounted for in Figure 2 of Gilbert et al. (1990, Study 1), assumed to reflect the not-seen responses of participants in that study. The pattern of responses unaccounted for in that study resemble those presented in the current research. However, without examination of the data it is unclear whether the supposition holds. Whiskers denote standard error.

There is a subtle but important difference between the Spinozan mind hypothesis and the forgetting effect suggested here. The Spinozan mind hypothesis predicts there is a tendency towards believing information is true in the early moments of comprehension. By contrast, the forgetting effect shows how the increased truth bias when rating interrupted propositions may reflect a tendency to forget false statements, thereby resulting in fewer false judgments and giving the appearance of a greater proportion of true judgments compared to false judgments.

Why might participants forget interrupted false statements more often than forgetting interrupted true statements? One possibility has been considered by a number of authors: true statements are more informative than false statements (Anderson et al., 2009, 2010; Glenberg et al., 1999; Hasson et al., 2005). Consider the proposition ‘a twyrin is a doctor’. When the statement is true there is a single clear meaning associated to the word twyrin: it is a doctor. But what if it were false – how might we interpret and encode the nonsense word twyrin when all that is known about it is that it does not mean ‘doctor’? The number of alternatives is vast. That the non-definitions of nonsense words, with little time given for encoding, may be forgotten is perhaps unsurprising.

Strategic ‘not-seen’ responding. There is another explanation aside from forgetting. Raters in the FT condition may not be forgetting, but rather using the not-seen response to indicate their uncertainty. That is, raters may not have truly forgotten the propositions, but only forgotten whether they were true or false definitions of the word. Instructions explicitly informed participants not to do so, but this by no means prevents participants from responding in this strategic fashion.

In discussing the three-way interaction above, focus has been cast upon the responses made in the forced choice FT condition. The three-way interaction also sheds light on the pattern of responding in the unforced FTU condition, and on the strategic use of the not-seen response. Judgments of interrupted true (M = .14, SD = .26) and false propositions (M = .14, SD = .26) were equally likely to result in relatively high level of not-seen responses, t (35) < 0.01, p > .999, d < 0.01. That is, interrupted propositions overall were likely to receive a high proportion of not-seen responses. Uninterrupted true (M = .04, SD = .15) and false (M = .08, SD = .13) propositions were also not significantly different from one another, t (35) = -1.14, p = .203, d = -0.29. Thus interrupted true and false items were no more likely to be remembered, and uninterrupted true and false items were no more likely to be remembered.

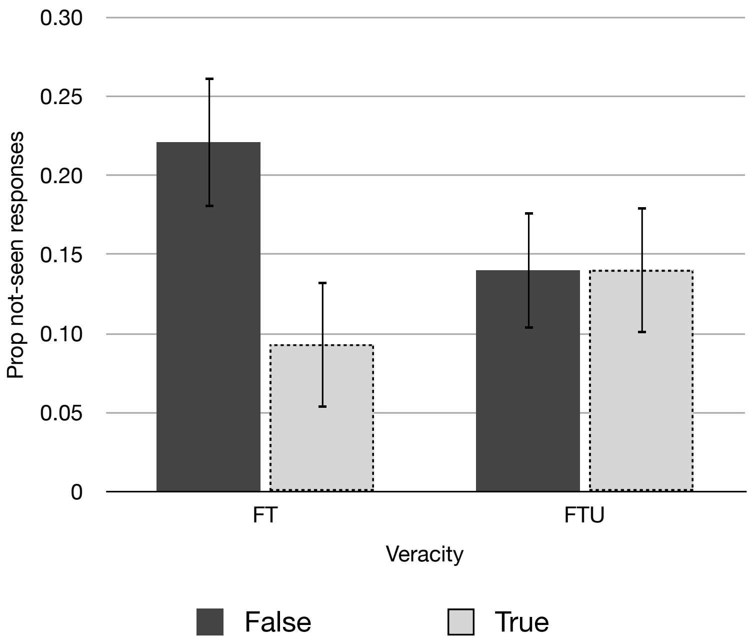

Figure 15. Not-seen responses when judging true and false propositions that were interrupted during learning. Whiskers denote standard error.

Consider the pattern of not-seen responding to interrupted true and false statements in both the FT and the FTU conditions (Figure 15). In the FT condition interrupted false statements received a higher proportion of not-seen responses than interrupted true statements. In the FTU unforced condition there was no significant difference in their pattern of not-seen responses to interrupted true and false statements. At first glance it could seem to suggest that the additional use of the not-seen response to false statements in the FT condition compared to the FTU condition is a strategic use of the response where those in the FTU condition would instead explicitly indicate their uncertainty. However, the non-significant difference in the use of the not-seen responses in the FTU condition is not evidence in support of the contention that there was truly no difference in how the FTU condition used the not-seen response. A Bayes factor was calculated in order to examine whether there truly was no difference, using a Cauchy prior with a scaling parameter $r = 0.5$, comparing a model with the Response Condition x Veracity interaction to a model without the interaction, and full-specified random effects on participants and stimulus proposition. Bayes factors near 1.0 indicate an inability of the data to speak to either the null or the alternative hypothesis due to a lack of significant power. A Bayes factor of 0.9 was calculated, failing to allow inferences regarding the interaction to be drawn.

In summary, a further analysis of the design of the experiment is not well positioned to test the Spinozan effect or the adaptive decision-maker account because the truth bias may reflect forgetting. The three-way interaction also hints at the possibility of a strategic use of the not-seen response, although statistical analyses determined there was a lack of sufficient power to determine whether this account could be supported.

Failure to Find a Reversed Spinozan Effect. The major prediction of this experiment was that in the forced choice condition a Spinozan effect would be observed: a truth bias should be seen when judging interrupted but not uninterrupted propositions. In the unforced choice condition it was predicted that the effect would reverse: external uncertainty due to interruption would result in uncertainty and, because they were not forced to judge, would indicate that uncertainty and show a reduced truth bias. However, internal uncertainty for uninterrupted propositions was expected to result in a truth bias: where uncertainty in this situation would, in forced choice conditions, be used as a heuristic towards believing the proposition false (see the discussion of Experiment 7 and the introduction to the current experiment), in unforced choice conditions they would no longer rely on this heuristic and so reduce their false responses, generating an artefactual truth bias. Whilst this latter prediction was upheld – raters were truth biased when rating uninterrupted propositions – there was also a truth bias when judging previously interrupted statements, contrary to predictions.

That there are differences in truth responding between the forced and unforced choice conditions leads to the natural question of how judgments of uncertainty were cast. This may shed light on the failure to support the predictions. A 2 (veracity) x 2 (interruption) ANOVA was conducted on the proportion of unsure responses in the FTU condition. That is, the sum of the unsure responses divided by sum of the unsure, true, false and not-seen responses. Surprisingly, all main effects and interactions were non-significant (all ps > .310), suggesting the use of the unsure response was not able to account for the differences between these conditions.

A second possibility is that participants in the FTU condition were forgetting interrupted statements less often than participants in the forced FT condition, and so do not experience the external uncertainty we predicted to be associated with a lack of bias. However, the data do not support this account. An independent samples *t*-test found no significant effect, $t\ (77) = 0.38$, $p = .706$, $d = 0.09$. A Bayes factor was calculated using a Cauchy prior with a scaling factor of r = 0.5, comparing a full model with fully specified random effects on participants and stimulus item to a simpler model without the response condition main effect. The data supported the null hypothesis of no difference: in order to accept the alternative hypothesis, prior odds favouring it of greater than 6 would be needed.

Finally, given that interruption appeared to result in similar degrees of forgetting for both the FT and FTU conditions, it may be the case that raters were more likely to forget interrupted false items than interrupted true items, regardless of response condition. As such, the unexpected truth bias in the FTU condition when rating interrupted statements may simply reflect a forgetting of interrupted false statements. The 'Strategic not-seen responding' subsection above explored the interaction between interrupted true and false statements with the response condition, and a Bayesian analysis determined there was not sufficient power to test this hypothesis. Here we are solely interested in whether interruption resulted in a higher degree of forgetting for interrupted false versus true statements, regardless of FT or FTU response condition. A paired-samples *t*-test confirmed raters were significantly more likely to forget false interrupted statements ($M = .18$, $SD = .27$) than true interrupted statements ($M = .11$, $SD = .24$), $t\ (78) = 2.17$, $p = .033$. This post-hoc exploratory analysis may suggest the truth bias observed in the FTU unforced response condition was found for interrupted statements because they were more likely to forget false items than true items.

However, given that the structure of this task is not suitable for testing the Spinozan effect, because the effects can be better explained in terms of forgetting than in terms of automatic believing, coupled with the exploratory and post-hoc nature of this examination, this explanation can not yet be accepted. Nonetheless, it offers some promise for an adaptive decision-maker account. Future research should return to a simpler design where the forgetting of items is less likely, such as the design used in Experiments 5 and 7.

Summary. A replication of Gilbert et al. (1990, Study 1) was carried out in the FT forced choice condition. The Spinozan effect was replicated: there was a truth bias for those statements previously interrupted, but a reduced biased for uninterrupted statements. However, on further exploration it was seen that forgetting could account for these differences: participants were more likely claim they had not seen interrupted false statements compared to true statements. As such, the structure of this task is unsuitable to test whether raters are Spinozan or whether, as claimed here, they are adaptive decision-makers.

It is possible that raters were not truly forgetting in the FT forced choice condition, but were using the not-seen response strategically to indicate their uncertainty. However, the lack of power did not permit an examination of this possibility.

It was surprising to note the unforced choice condition showed a biased pattern of responding when interrupted, counter to predictions. The pattern of uncertainty responses was unable to provide any insight into this bias. Whilst it was not possible to test the interaction between the response conditions and the true and false statements that were interrupted during learning, due to a lack of sufficient power, a post-hoc *t*-test found raters overall, regardless of response condition, were more likely to forget false statements that were interrupted during learning than they were true statements that were previously interrupted. As such, the truth bias for interrupted items, regardless of FT or FTU response condition, may be seen to reflect a greater forgetting of false items than true items and thus result in an artefactual truth bias.

General Discussion: Chapter 5

The HAM dual-process account received little support in Chapter 4. Although it has been claimed that the time scale of minutes may be reflect a shift from heuristic to analytical processing (Masip, Garrido, et al., 2009; Masip et al., 2010), the timings may be too coarse to capture the two processes in operation. The Spinozan mind theory (Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993) proposes a default-interventionist model that operates during the act of comprehension. It claims that belief is a default state that is attributed to all incoming information: it is a necessary component of comprehension. It is only afterwards that a revision can take place. The account has had an impact across a range of research areas (Burge, 1993; Fitzsimons & Shiv, 2001; Hasson et al., 2005; Millar & Millar, 1997; Pennycook et al., 2012).

An alternative account, an informed Cartesian account, can also explain these findings, if one allows for the possibility that contextual knowledge can be brought to bear on the decision. As discussed in Chapter 2, a number of studies have begun challenging the experimental data supporting the Spinozan view, showing how processing times for false statements may be an artefact of the meaningfulness of the true and false propositions used in prior research (Anderson et al., 2009; Fraundorf et al., 2013; Glenberg et al., 1999; Hasson et al., 2005; Mayo et al., 2004). Here it is claimed that the Spinozan effect is actually an adaptive response to the unfortunate position of being forced into judgment before a decision has been reached. When pushed to respond in the absence of evidence, we can rely on prior knowledge. Experiment 5 showed how the act of forcing participants into making a judgment resulted in an early bias towards believing, whereas those not forced to make a judgment showed no such early bias.

Chapter 6: Social Orientation Theory

The truth bias extends across various domains of research, from findings in psycholinguistics showing a processing advantage for affirmed over negated sentences (Carpenter & Just, 1975; Clark & Chase, 1972; Clark & Clark, 1977; Mayo et al., 2004; Trabasso et al., 1971; Zwaan et al., 2002) to research in belief formation showing an increased tendency to believe information when it is coupled with related but uninformative additional details (Fenn, Newman, Pezdek & Garry, 2013; Newman, Garry, Bernstein, Kanter & Lindsay, 2012). That the truth bias holds such range has led this thesis to begin from a bottom-up approach, examining first whether cues in the environment can account for the truth bias (Experiment 1) and from there to the cognitive processes underlying the judgment process (Experiments 2-8).

Yet our subject has been one that at its base is a social task, one that requires inferences of others' intent. The intent of this thesis is to explore decision-making processes in uncertain socially situated environments. This chapter takes an initial exploration of the social influences that guide the belief formation process.

Although the truth bias places limits on accuracy rates, it may serve a useful and adaptive social purpose (DePaulo et al., 1996; Lewis, 1993; Vrij, 2008; Vrij, Granhag & Porter, 2010). Placing trust and our belief in others is important for maintaining relationships: holding a sceptic's viewpoint and challenging each point that raises suspicion is soon to lead to difficulties (Bell & DePaulo, 1996; Clark & Lemay, 2010; Cole, 2001; DePaulo & Bell, 1996; Guthrie & Kunkel, 2013; Miller et al., 1986).

What is more, in actively seeking out the truth we may come to discover unpleasant facts: that the thoughtful birthday gift you sent was perhaps not as well received as you would have hoped it to be: sometimes we wish not to know the truth and do not seek it out, known as the Ostrich effect (DePaulo et al., 2003; Ekman, 1992; Vrij, 2008; Vrij, Granhag & Porter, 2010).

For these reasons we may bring a degree of willing naïveté to the lie detection table. Avoiding the socially aggressive act of accusation and the potentially aggressive response to the challenge of their honesty may outweigh the benefits of discovering the truth. The accusatory reluctance account (DePaulo & Rosenthal, 1979; Ekman, 1992; Miller et al., 1986; O'Sullivan, 2003; O'Sullivan et al., 1988; Vrij, 2008; Vrij, Granhag & Porter, 2010) makes just such a claim, and suggests the truth bias results from an unwillingness to challenge others because of the implicit social rules that govern interaction.

In support of a role for social information in the truth bias, it has been shown people are less likely to give negative feedback to those with whom they are in close relationships with compared to more distant relational partners (DePaulo & Bell, 1996; DePaulo & Kashy, 1998; Lemay & Clark, 2008; Uysal & Oner-Ozkan, 2007; see also Boon & McLeod, 2001), and when it is given it is often subtle (Clark & Lemay, 2010; Metts, 1989; Swann et al., 1992). As the relationship becomes closer and the length of the relationship increases, there is an increased tendency toward believing (Argyle & Henderson, 1984; McCornack & Levine, 1990; McCornack & Parks, 1986; Stiff et al., 1992; Stiff et al., 1989; van Swol et al., 2012; see also Boyes & Fletcher, 2007) and trusting them (Gaertner & Dovidio, 2000; Voci, 2006).

Whilst there is research consistent with an accusatory reluctance account, there has been as yet no direct empirical test of its fundamental claim: whether the truth bias stems from the perception of another person as a social being, one that can feel aggressed upon and in turn be aggressive. Mere implied social presence, as is the case in videotaped statements used in most lie detection experiments, must be sufficient to invoke an accusatory reluctance if it is to explain the truth bias that is so often observed in response to videotaped interviews.

This chapter opens with a consideration of the minimal conditions under which a truth bias may be expected, given a social context effect. Experiment 9 empirically tests the minimal conditions for an accusatory reluctance account of the truth bias. Because at its core the account requires that mere social presence is sufficient to invoke a truth bias, other accounts of the truth bias as a willing sacrifice of accuracy to conform to social rules and norms are implicitly tested. If the truth bias reflects socialisation practices, other social-based accounts must also meet these minimal conditions. This experiment finds that implied social presence is not a sufficient condition to invoke a truth bias, failing to support a social account.

Experiment 9: Socialisation Practices

This experiment seeks to address the question of whether mere implied social context is sufficient to invoke a truth bias. Because this study seeks out the minimal conditions, all other potentially relevant information in the behaviour that may also guide biased responding is removed. Participants were given stick-figure models built from body motion capture recordings with no verbal or paraverbal information presented. They were led to believe these recordings were either recordings of human speakers (social context condition) or computer-generated animations (non-social context condition). A lack of bias was expected when they believed the speakers were actually computer animations, but that the truth bias would be observed under conditions of an implied social context, i.e. when the speakers were believed to be human.

Method

Materials. In order to test the influence of social presence, a new stimulus set had to be generated. Twelve speakers (7 female, age M = 29.58, SD = 8.85, range 20 to 47 years) signed up to take part in an 'eyewitness' study supposedly exploring the links between eyewitness testimony and distracting clothing. Participants were sampled from a paid participant pool and were paid £5 for participation. In all studies reported participant signed informed consent and were given the option to withdraw retrospectively. Participants donned a sports top, baseball cap and loose fitting tracksuit bottoms, supposedly in order to remove any effects the style of clothing may have on our officers' lie-truth judgments. The clothing had sewn into it reflective markers that allowed us to capture bodily motion. The sports top had 15 reflective markers sewn into it, the hat 5 markers and the tracksuit bottoms had no markers attached. In addition, one reflective marker was added to the back of each of the hands of the participant, resulting in a total of 22 reflective markers (see Figure 16).

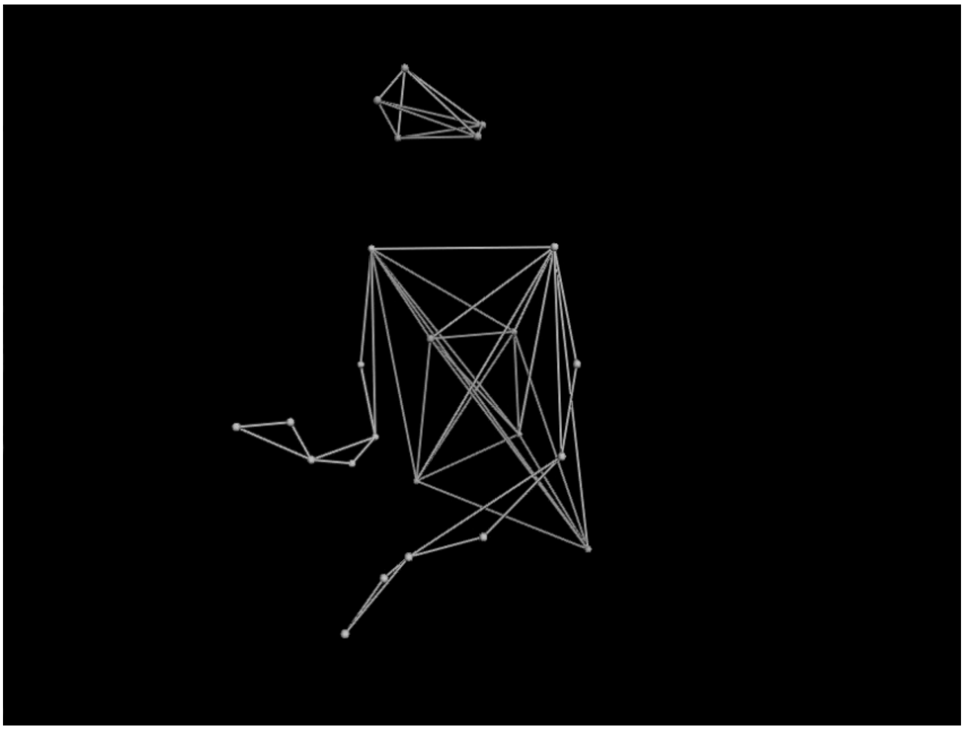

Figure 16. A scene taken from a stimulus item presented to participants in Experiments 9 and 10. There was no audio accompanying these videos.

Four stacks of three envelopes were offered to participants, supposedly to randomise which mock crime video they would see (in fact there was only one video). The first envelope told them they were about to watch a mock crime and were asked to remember as much as they could. The video presented two pieces of video footage recorded from a high angle with an imprint of the street address and time stamp (but in colour and presented at 30 frames per second) to imitate CCTV footage (Figure 17). The two pieces of footage were presented side-by-side, as might be the case of security footage from two cameras viewing the same scene from different angles.

Figure 17. Scene taken from the CCTV footage shown to participants who made up the stimulus set.

The victim in the mock crime was seen to hug a girl before walking onto the street corner with a suitcase whilst talking on a mobile phone. The perpetrator was seen talking with another person. As the victim approached the street corner the perpetrator tried to steal the phone but failed. A fight broke out between the two, and the perpetrator threw the victim's suitcase at the victim. The victim's friend then returned and attempted to pull the perpetrator off the victim, but was pushed away, at which point she, the friend, left the scene. The perpetrator's friend joined the fight and eventually they succeeded in stealing the phone. The victim never showed any aggression nor used violence.

After watching the video, participants were given the second envelope from their freely chosen stack that told them they were the sole witness to this crime and that they should act as though the video footage did not exist. It further said they were to be interviewed by a Metropolitan police officer who had not seen the crime footage or the contents of the envelopes, and that he would be attempting to determine whether the participant was lying or telling the truth. The envelope either instructed them to lie and protect the perpetrator or to tell the truth and defend the victim. They were told if they were successful in persuading the police officer they were being honest they would receive an additional cash bonus at the end of the experiment.

After three minutes of solitary preparation time, participants were introduced to and interviewed by the posing police officer. They answered six questions pertaining to the details of the crime (e.g. “what happened with the suitcase?”) and then were asked to give a full recall of the set of events from the moment the people entered the crime scene until everyone present had left. This was then repeated so that the third envelope encouraged lying or truth telling, depending on whether they had told the truth or lied in the previous interview, respectively. They were then introduced to and interviewed by a second posing officer, also male.

At the end of the experiment all participants were informed of the true aim of the experiment and given the opportunity to withdraw retrospectively, of which none chose to do so.

Body motion data was recorded during the interviews by six Vicon MX-F40 cameras connected to a Vicon MX Giganet core unit via Gigabit Ethernet. Data was captured at 200 frames per second with a 4-megapixel resolution (2352H x 1728V) by Vicon Nexus software, capable of recording with an accuracy of 0.1mm. The body motion data was extracted offline resulting in a wire model shown in Figure 16. The first 30 s of each speaker’s deceptive and honest full recall statements was taken to provide a total of 24 statements (12 honest, 12 deceptive) of equal length, forming our video stimuli. These were split into two video sets consisting of 6 truthful and 6 deceptive statements, such that participants who rated the stimulus set saw each speaker only once, viewing either that person’s lie or their truth. The body motion capture did not collect audio information, thus the only information available to the raters was a video of the wire frame model.

Participants. Ninety-one undergraduate students took part in a lie detection study either voluntarily for no reimbursement or as part of their course fulfilment. Two participants were removed because their response times were larger than three standard deviations from the mean, suggesting they were not paying attention to the task. This left 89 participants, 74 of which were female (age $M = 19.51$, $SD = 1.31$, range 18 to 24 years).

Procedure. After signing consent, an instruction screen informed participants they were to view videos of wire frame models. In the human belief condition, participants were led to believe the wire frames were captured from people undergoing an interview, as was indeed true. In the simulation belief condition, participants were led to believe the wire frames were computer generated and thus were not videos of real people. Specifically, participants in the human belief condition were told, ‘The clips you will watch in a moment are movements of people giving true and false testimony’, whilst those in the simulation belief condition were told, ‘This study aims to test a developed simulation model. The clips you will watch in a moment are generated by a computer that imitates the movements of people giving true and false testimony.’ They were further told videos were sampled randomly and so they could see more lies than truths, more truths than lies or an equal split. They would need to judge whether the video was thought to be (a simulation of) a lie or (a simulation of) a truth.

After two practices, the instructions were presented again on screen before the 12 experimental trials began. One of the two video sets was selected randomly and presented in a randomised order.

Design. Raters were assigned randomly to either the human belief condition ($n = 42$) or the simulation belief condition ($n = 47$). The veracity of the speaker's statement (lying or telling the truth, within subjects) and the belief condition (human or simulation belief, between subjects) served as the independent variables, and the PTJ, accuracy scores, and signal detection measures served as the dependent variables.

Results

The accusatory reluctance account would predict a higher proportion of 'truth' judgments when participants believed they were rating real people than when they thought they were rating computer simulations. We did not find support for this account. A Bayes factor indicated the data shift the relative plausibility of the model that included the experimental manipulation towards the model that did not, providing evidence in favour of the null effect of the human belief manipulation.

A 2 (veracity: lie, truth, within subjects) x 2 (belief condition: human, simulation, between subjects) mixed ANOVA was conducted first on the PTJ. A PTJ above 0.50 indicates a bias towards judging videos as truths. The human belief condition appeared to show little bias when rating lies ($M = .52$, $SD = .19$) or truths ($M = .51$, $SD = .23$). The simulation belief condition showed some tendency towards a truth bias when watching lies ($M = .55$, $SD = .19$) and less so when watching truthful statements ($M = .52$, $SD = .23$), with the means appearing to show little difference between the human and simulation belief conditions. Indeed, and contrary to the predictions of the accusatory reluctance account, the ANOVA confirmed there was no statistically significant effect of belief condition, F (1, 87) = 0.68, p = .413, η^2 = 0.01, nor any interaction with veracity, F (1, 87) = 0.08, p = .778, η^2 < 0.01. In addition, there was no main effect of veracity, F (1, 87) = 0.39, p = .537, η^2 < 0.01.

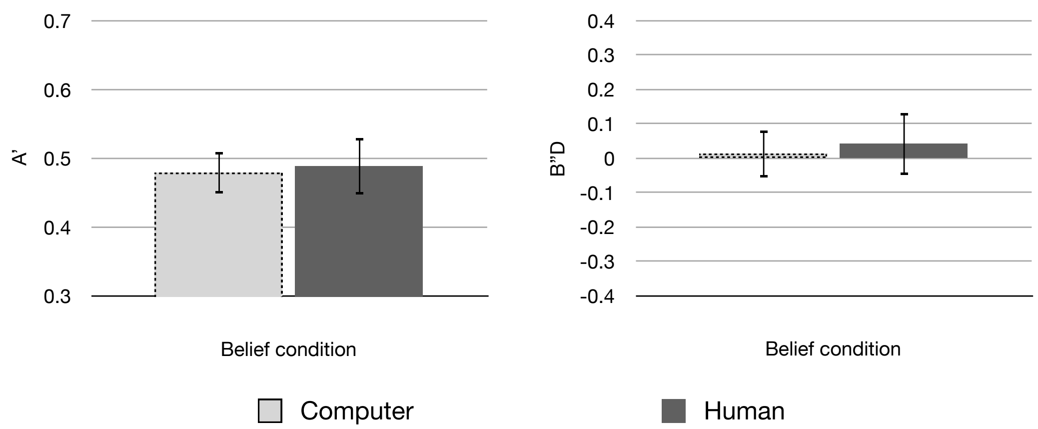

Figure 18. (a) Signal detection measure of accuracy (left) and (b) response bias (right) when the speakers were portrayed as either computer generated or human speakers. Whiskers denote standard error.

The same ANOVA conducted on accuracy scores found no main effect of veracity, F (1, 87) = 3.25, p = .075, η^2 = 0.04, belief condition, F (1, 87) = 0.01, p = .778, η^2 < 0.01, or Veracity x Belief Condition interaction, F (1, 87) = 0.68, p = .413, η^2 = 0.01.

To explore the independent effects of accuracy and response bias, signal detection measures of A' and B''_D were calculated. An independent samples *t*-tests compared the human and simulation belief conditions on their A' (M = .49 and .48, *SD* = .26 and .20, respectively) and B''_D scores (*M* = .04 and .01, *SD* = .59 and .47, respectively). Both A', *t* (87) = -0.19, *p* = .854, *d* = 0.04, and B''_D scores, *t* (87) = -0.26, *p* = .794, *d* = 0.06, did not show a statistically significant difference between the human and simulation belief conditions.

It was also considered whether raters exhibited accuracy above chance as well as a significant response bias in each condition independently. The accusatory reluctance claim as outlined here would predict no response bias observed in response to the (believed) computer-simulated movements but would exhibit a truth bias in response to the wire frames when they were believed to be portraying humans. However, neither the human condition, *t* (41) = 0.47, *p* = .639, *d* = 0.07, nor the simulation belief condition, *t* (46) = 0.20, *p* = .841, *d* = 0.02, showed a statistically significant response bias. In addition, both the human and simulation belief groups showed accuracy rates not significantly different from chance, *t* (41) = -0.27, *p* = .787, *d* = -0.04, and *t* (46) = -0.69, *p* = .496, *d* = -0.10, respectively.

A Bayes factor was calculated to gain more information about the null effect found above. A Cauchy prior distribution with r = 0.5 was specified on the effect size, where r is a scaling parameter (Jeffreys, 1961; Zellner & Siow, 1980). A complex model entered the belief condition and the veracity of the speaker as fixed effects, with the PTJ acting as the outcome variable. Fully specified random effects were included for rating participants and for speakers. This complex model was compared to a simpler model with the belief condition main effect removed. A Bayes factor indicated that in order to prefer the more complex model, we would need prior odds favouring it of greater than 12. This was taken as strong support for the null hypothesis of no predictive effect of belief condition.

Discussion

Calling someone a liar is an aggressive social act. A reluctance to break with social rules and accuse others of lying (O'Sullivan, 2003; see also DePaulo & Rosenthal, 1979) may account for the fact that we tend to believe what others are saying more often than their actual rate of honesty warrants (Bond & DePaulo, 2006). No truth bias was observed when the social elements of the stimuli were removed, consistent with the accusatory reluctance account. There are many differences between the wire frame stimuli in our simulation belief condition that found no truth bias and the full audio-visual recordings used in prior studies that show a truth bias (Bond & DePaulo, 2006; Zuckerman et al., 1979; Zuckerman, Kernis, Driver & Koestner, 1984). For instance, information about emotion and cognitive load may be observable in the face, and verbal cues may lend more credence to the speaker's tale. By itself this finding is consistent with the accusatory reluctance position, but it could also be accounted for by the relative paucity of information in the stimuli compared to typical lie detection tasks. Thus to test the accusatory reluctance account it is necessary to compare ratings of the same stimuli when raters believe they are social compared with non-social beings. It is claimed the social rules and practices governing interpersonal relationships cause the truth bias, and as such merely perceiving the stimuli as social beings should be sufficient to invoke a truth bias. However, no support was found for this claim: raters were equally unbiased when they believed the stimuli were humans as when they believed the stimuli were computer generated simulations. Also, raters who believed they were rating computer animations did not show greater accuracy rates, contrary to what may have been predicted (O'Sullivan, 2003).

This result was surprising; in prior research reducing the psychological distance between the rater and the speaker appears to influence the degree of truth bias such that those perceived closer benefit from an exaggerated truth bias whilst those who are being judged for their criminal intent suffer at the hands of a lie or investigator bias (McCornack & Levine, 1990; Meissner & Kassin, 2002; Stiff et al., 1989). Also, people report feeling more uncomfortable when lying to those to whom they feel closer (DePaulo & Kashy, 1998).

If the truth bias is not based in our understanding of the social world, it seems surprising that changes in the social relations between people influences the degree of truth bias. One possibility is that the use of social information is used in a more cognitive fashion, as another piece of evidence that contributes to the decision making process. We may not be adapting our behaviour to match the social expectations of our interacting partners and willingly sacrificing accurate decision-making. Instead, the information about how psychologically close others are to us can be taken as a cue to how much we ought trust others, for example (see O'Sullivan, 2003; Stiff et al., 1992; Wickham, 2013). Experiment 10 takes an initial step towards testing this claim.

Experiment 10: Social Relatedness as a Heuristic

To show social contextual information can be used in the decision-making process, it is important to account for other factors that tend to covary with relatedness. For instance, those who are close with one another are likely to share similar preferences (Allen & Wilder, 1979; Diehl, 1989; Tajfel et al., 1971) and an understanding of social rules built between them (Vrij, 2008). The minimal group paradigm allows for a separation between perceived relatedness and these other factors that, in prior research on lie detection, have been confounded (e.g., McCornack & Levine, 1990; McCornack & Parks, 1986).

The minimal group paradigm (Tajfel, 1970; Tajfel et al., 1971) randomly assigns participants into arbitrary groups that have no meaningful relationship with the task set them. Because assignment to groups is random, the only link between participants within a given group and the only difference between the ingroup and the outgroup is that they have been randomly grouped.

The current experiment makes use of a classic paradigm in this area Tajfel et al. (1971, Study 2). Participants were supposedly grouped based on their preference for one of two artists, but were actually assigned randomly to one or other group. Studies such as this one show that despite there being no cost to rewarding both ingroup and outgroup members equally, ingroup members are favoured (see also Brewer & Silver, 1978). The same effect is found even when participants were aware they were randomly grouped (Billig & Tajfel, 1973).

Intergroup biases reliably surface using the minimal social grouping paradigm, which continues to be a practical research tool decades after its conception (e.g., Chen & Li, 2009; Leonardelli & Brewer, 2001; Postmes, Spears & Lea, 2002; Richardson et al., 2012; Worchel, Rothgerber, Day, Hart & Butemeyer, 1998). In three reviews of intergroup biases in minimal social groups, Brewer (1979, 1999) and Otten and Mummendey (2000) concluded there was an asymmetry between ingroup preference and outgroup derogation. Group members more often favour the ingroup rather than discriminate against the outgroup.

In line with these reviews, a truth bias was predicted to be found when rating ingroup members, but not when rating outgroup members, who could potentially succumb to derogation and thus a lie bias.

The weaker version of the accusatory reluctance account was also tested, as per Experiment 9, using accuracy rates. In line with the accusatory reluctance account and the findings of O'Sullivan (2003), lower accuracy was expected when rating those who are psychologically closer to the rater because the reluctance to publicly make (accurate) lie judgments should be increased. This is considered a weaker test of the account because increased accuracy is mediated via a reduced truth bias. A direct examination of the truth bias allows for a stronger test of the account.

Methods

Materials. Twelve digital reproductions of abstract paintings by the artists Paul Klee and Wassily Kandinsky that were judged by the experimenter to be not too dissimilar from each other were used for collecting participants' art preferences. The wire frame video stimulus set from Experiment 9 was used here also.

Participants. Sixty undergraduate students participated in part fulfilment of their studies, 46 of which were female (age M = 20.02, SD = 2.90, range 18 to 40 years).

Procedure. Participants were first told they would see a range of artwork produced by modern artists, and that they should select the piece they most preferred. They were told to go with what they felt instinctively was the correct choice. After reading this instruction, a pair of images was shown side-by-side and the participant clicked on the image they preferred. After 12 such trials, the computer randomly determined whether they were in the stars or moons group supposedly based on their preferences. The text ‘You are in the stars/moons group’ was displayed for 3 s above an image of either a red star or a blue moon, accompanied with a purple circle in the background.

The rating instructions were then presented, this time with their group icon placed in the top-left corner of the instruction screen. Participants in all groups saw the same instruction as given to the human belief condition in Experiment 9. They were additionally told the speakers in the stimulus set had completed the same art preference-rating task and had also been assigned to a group in the same manner the participant had just been assigned.

Prior to trial onset, participants were shown two icons on screen: their own group icon on the left, and the group icon of the speaker they were about to view on the right. Below this read ‘You are in the Stars/Moons group. The next person is in the Stars/Moons group.’ Above the left icon the word ‘You’ was displayed, and above the right icon the words ‘Next Person’ were displayed. This remained on screen for 3 s before the trial began. Otherwise a trial proceeded as in Experiment 9, with a lie-truth rating collected at the end of each video. After two practice trials, the instructions were presented again and the remaining 12 experimental trials followed.

Design. Participants were assigned to the ‘moons’ (n = 31) or the ‘stars’ (n = 29) group at random. The grouping is unimportant for this study; what is important is whether the speaker was depicted as an ingroup member, i.e. as also a member of the moons/stars group, or as an outgroup member. Group membership (ingroup, outgroup) was a within-subjects independent variable. Six of the statements (three honest, three deceptive) were randomly presented as ingroup members whilst the remaining six statements (three honest, three deceptive) were presented as outgroup members. The dependent variables were the PTJ, accuracy, and the A' and B''_D scores.

Results

In accord with the hypothesis, an ingroup preference was observed. Ratings of ingroup members were truth-biased whilst ratings of outgroup members were not significantly biased. There were no significant differences in accuracy when rating in and outgroup members, but ingroup members were detected at above chance rates.

A 2 (veracity: lie, truth) x 2 (intergroup relationship: ingroup, outgroup) within-subjects ANOVA was conducted on the PTJ. A marginally significant main effect of veracity was found, $F(1, 59) = 3.60$, $p = .063$, $\eta^2 = 0.06$. Ratings of truthful statements were more often rated as truths ($M = .58$, $SD = .21$) than were deceptive statements ($M = .49$, $SD = .26$). There was also a marginally significant main effect of intergroup relationship, $F(1, 50) = 3.68$, $p = .060$, $\eta^2 = 0.06$. Speakers in the same group as the rater received a higher PTJ ($M = .57$, $SD = .19$) than speakers in the outgroup ($M = .51$, $SD = .21$).

The same ANOVA conducted on the accuracy scores found a marginally significant main effect for veracity, $F(1, 59) = 2.80$, $p = .056$, $n = 0.06$. Truthful behaviours ($M = .58$, $SD = .21$) were more accurately judged than deceptive behaviours ($M = .51$, $SD = .21$). There was also a marginally significant interaction between intergroup relationship and veracity, $F(1, 59) = 3.68$, $p = .060$, n = 0.06. Bonferroni-corrected post-hoc *t*-tests found ingroup truths ($M = .64$, $SD = .32$) were detected more accurately than lies ($M = .50$, $SD = .32$), $t(59) = -2.90$, $p = .006$, $d =$

0.44, whilst accuracy in detecting outgroup members was not significantly different when judging truths (*M* = .52, *SD* = .33) or lies (*M* = .53, *SD* = .26), *t* (59) = -0.03, *p* = .819, *d* < 0.01.

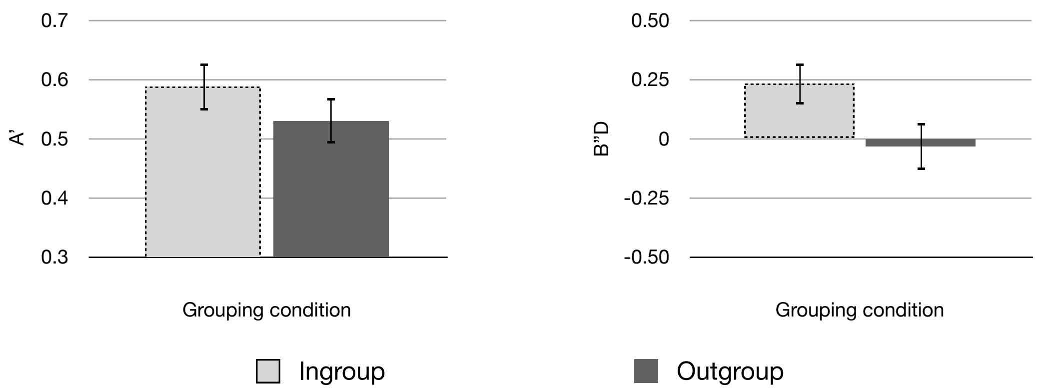

Figure 19. (a) Signal detection measures of accuracy (left), and (b) response bias when the speakers were portrayed as either ingroup or outgroup members (right). Whiskers denote standard error.

A paired samples *t*-tests on the B"$_D$ response bias scores showed a significant effect, *t* (59) = -2.01, *p* = .049, *d* = -0.36. Raters were truth biased when judging ingroup members (*M* = -0.23, *SD* = 0.68) whilst, if anything, somewhat more inclined to be lie biased when rating outgroup members (*M* = 0.03, *SD* = 0.77). One-sample *t*-tests confirmed ratings of ingroup members were significantly truth biased, *t* (59) = -2.64, *p* = .011, *d* = -0.34, but were not significantly biased when rating outgroup members, *t* (59) = 0.33, *p* = .744, *d* = 0.04.

The A' accuracy scores when rating ingroup (*M* = .59, *SD* = .31) and outgroup speakers (*M* = .53, *SD* = .30) did not differ significantly, *t* (59) = 1.23, *p* = .223, *d* =

0.20. In addition to being truth biased, raters achieved above chance accuracy when rating ingroup members, t (59) = 2.18, p = .033, d = 0.29, but were no more accurate than chance when rating outgroup members, t (59) = 0.77, p = .445, d = 0.10.

Discussion

The most minimal form of group membership leads to greater trusting of ingroup members (Gaertner & Dovidio, 2000; Voci, 2006). More realistic but still random and arbitrary groupings can even lead to threats of violence with long-term consequences for the interactions of individuals (Sherif, Harvey, White, Hood & Sherif, 1954/1961). It was anticipated raters would consider ingroup members to be more likely to tell the truth compared to outgroup members. The findings supported the prediction. Further analysis showed ratings of ingroup members were truth biased, whilst ratings of outgroup members showed no statistically significant bias in their judgments. That is, there were ingroup benefits (from the speakers' perspective) insofar as there was a truth bias but not outgroup derogation, consistent with previous intergroup relations research (Brewer, 1979, 1999; Otten & Mummendey, 2000).

No significant difference in accuracy was observed when rating ingroup versus outgroup members. However, and in direct contrast to an accusatory reluctance account, whilst accuracy in judging outgroup members was at chance rate, accuracy in judging ingroup members exceeded chance. It is important to note the analyses conducted on the accuracy and the PTJ scores confound bias with accuracy. It is not clear whether accuracy in judging ingroup members' truths was greater than in judging their lies because they were more biased towards saying truth or because they more readily noticed cues to honesty. Although signal detection measures can separate out the effects of bias and accuracy, they do not, by necessity, allow for a distinction in the accuracy in judging truthful and deceptive behaviours. As such, is it not possible to conclude raters were necessarily more accurate in judging truthful than deceptive ingroup behaviours as exhibited in the raw accuracy scores because this difference may be accounted for by differences in bias. Nonetheless, it is interesting that ingroup members' behaviours were overall accurately judged above chance rates whilst outgroup members' behaviours were not. Recall an accusatory reluctance account suggests outgroup members should be *more* accurately judged than ingroup members because of the reduced reliance on biased judgments that maintain harmony.

One possible explanation of this effect stems from social identity theory (Tajfel, 1981; Tajfel & Turner, 1979). Self-identity is tied up in group belonging. Individuals may form their self-image based around characteristics of their group (Otten & Epstude, 2006; Smith & Henry, 1996), which may account for the preferential treatment ingroup members are given (Allport, 1954; Brown, 2000; Grant & Brown, 1995). There may be a greater motivation to detect ingroup members' truthful behaviours in order to reaffirm the moral constitution of the group and, by extension, of the self. Overall accuracy rates may be driven by accuracy in detecting truths more than lies, in line with prior research (Levine et al., 1999). An alternative cognitive account rests on findings that show information provided by ingroup members is processed more deeply, better remembered, and remembered in greater detail than information provided by outgroup members (Sporer, 2001; van Bavel, Packer & Cunningham, 2008). Additional processing effort may account for the above-chance accuracy rates when judging ingroup members.

There are limitations in interpreting the accuracy findings. Consideration should be given to the lack of a statistically significant difference in the current experiment between the accuracy in judging ingroup compared to outgroup members. Further empirical exploration should be carried out to replicate the accuracy findings reported here with consideration giving to these self-identity and cognitive accounts.

There are further limitations to interpreting the current findings. The artificiality of the current task (see Aschenbrenner & Schaefer, 1980), using body motion capture recordings, was used in order to closely replicate the setup of Experiment 9, an experiment that required minimal information available to the rater. In addition, the minimal context served to reduce other factors that are confounded with relatedness, such as familiarity with the speaker. However, it should be noted that ingroup biases have a greater magnitude in more consequential group relations (Mullen, Brown & Smith, 1992; Riek, Mania & Gaertner, 2006; Scheepers, Spears, Doosje & Manstead, 2006). Under more realistic settings stronger ingroup preferences may be expected. Nonetheless, this speculation remains to be tested.

It is also unclear whether the raters used the ingroup-outgroup information as a simple heuristic or whether raters felt closer to those with whom they were paired. In the latter case, participants may have made inferences beyond the presented information. For example, in two diary studies, DePaulo and Kashy (1998) found people told fewer lies to people they felt close to than to strangers and acquaintances, although it is worth noting that 92% of respondents in one study admitted lying to an intimate partner (Cole, 2001) and in another study people lied in nearly one in 10 interactions with an intimate partner (DePaulo & Kashy, 1998). Because the participants and the speakers had supposedly been grouped on their preferences, raters may have felt closer to the individuals and believed they were less likely to lie.

Whilst this proves to be a serious drawback for the current study, it is important to consider that the effects of the minimal group paradigm have been shown to hold even when raters know they have been randomly assigned to groups (Billig & Tajfel, 1973). Unfortunately, members of our ingroup are ascribed qualities that we perceive as true of ourselves (Otten & Epstude, 2006; Smith & Henry, 1996), regardless of the initial dimensions on which grouping took place (Allport, 1954; Billig & Tajfel, 1973; Tajfel et al., 1971). Thus merely being aware of the random grouping factor may not be sufficient to avoid this potential confound.

Although these drawbacks place limitations on what can be drawn from this study, this is the first attempt to empirically assess how the degree of relatedness influences decision making independent of other potentially confounding factors such as greater experience with those who we are closer to. Whether beliefs about social relatedness lead to an attribution of a range of positive characteristics we attribute to ourselves or whether it is used as a more simplified social heuristic (‘people like me tell the truth’) remains to be seen. However, it is clear that social relatedness does influence the degree to which we believe others, independent of familiarity with the speaker, prior experiences with that individual, and so on.

General Discussion: Chapter 6

According to the accusatory reluctance account (O'Sullivan, 2003; see also DePaulo & Rosenthal, 1979), labelling another person a liar is a socially aggressive act that can have negative, potentially aggressive repercussions. Thus rather than challenge the implied social rules, a willing naïveté may be preferable (DePaulo et al., 2003; Ekman, 1992; Vrij, 2008; Vrij, Granhag & Porter, 2010). After all, a relentless distrust in others is liable to damage our relationships with them (Clark & Lemay, 2010; Miller et al., 1986). Fundamental to the accusatory reluctance account is the tendency to believe others 'merely' because they are social agents, and as such binds raters to the social rules of politeness in judging others. Whilst this account is cited as one potential explanation of the truth bias, it has never been directly tested. Experiment 9 represents the first examination. By manipulating whether the behaviours presented were believed to be of another (human) social agent or of a computer-generated animation, it was possible to test whether the implied social contract between rater and speaker was sufficient to result in a truth bias. No support was found for this account: there was similarly unbiased responding when the behaviour depicted was believed to be either human or computer generated. In two experiments there was no evidence for the position that accuracy increases when speakers are seen as more psychologically distant from oneself, as has been previously suggested (O'Sullivan, 2003).

That raters report being uncomfortable when viewing deceptive behaviour (Anderson et al., 2002; DePaulo et al., 1997; Toris & DePaulo, 1985) was considered as suggestive evidence in favour of the accusatory reluctance account. Discomfort may be associated with the consideration of violating the implied trust social contract of trust and aggressively challenging the speaker. This discomfort, the present findings would suggest, is not caused by an apprehension. One alternative is that the sorts of behaviours liars display cause raters to feel uncomfortable; perhaps raters are empathising with the discomfort liars may be experiencing. Whatever the explanation, it is concluded only that accusatory reluctance appears not to account for the truth bias and so discomfort may be caused by other factors.

Making Use of Social Information

On the face of it, Experiment 9 appears to contradict previous findings. Consider that the greater the length of an intimate relationship, the greater the tendency to believe the speaker (McCornack & Levine, 1990; Stiff et al., 1989; see also Boyes & Fletcher, 2007). Similarly, the psychologically more distant relationship between a police investigator and an interviewee exhibits a complimentary bias towards disbelieving the speaker (Kassin et al., 2005; Masip et al., 2005; Meissner & Kassin, 2002). That social presence is not sufficient to cause a truth bias led to the consideration that information about the nature of the relationship may be used to guide the decision.

Again, utilising a minimal social context, Experiment 10 manipulated the perception of the relationship between rater and speaker independent of other potentials cues, such that raters believed they were making intragroup (ingroup) or intergroup (outgroup) judgments. Previous research has shown greater trust is placed in ingroup compared with outgroup members (Gaertner & Dovidio, 2000; Voci, 2006). We found when rating outgroup behaviour there was no bias towards either believing or disbelieving the speaker, similar to ratings made of perceived computer-generated and human-generated behaviours in Experiment 9. In line with previous research on intergroup perception (Brewer, 1979, 1999; Otten & Mummendey, 2000), even under the most minimal of social contexts there was an observable truth bias when judging ingroup members. That is, ingroup speakers were judged preferentially with no outgroup derogation.

Early findings stemming from the minimal group paradigm led to the formation of social identity theory (Diehl, 1990). Social identity theory (Tajfel, 1981; Tajfel & Turner, 1979) proposes one's identity is defined by the characteristics that form the basis of one's ingroup. As such, members of the ingroup are ascribed qualities that individuals perceive as true of themselves (Otten & Epstude, 2006; Smith & Henry, 1996), regardless of the initial dimensions on which grouping took place (Allport, 1954; Billig & Tajfel, 1973; Tajfel et al., 1971). By exhibiting belief in ingroup members, raters may also be reaffirming their own identity as an honest person (Allport, 1954; Brown, 2000; Grant & Brown, 1995). Extending what is known about the self, a vast array of characteristics, to others based on relatively sparse information (grouping information alone) may typically prove to be an effective policy. However, what is not clear from these initial studies is whether people are using such a social heuristic (Fan et al., 1995; Fiedler & Walka, 1993; O'Sullivan et al., 1988; Zuckerman, Koestner, et al., 1984) or whether there are more complex processes of representing others in a functionally equivalent way to the ways in which we represent ourselves (see Sebanz, Bekkering & Knoblich, 2006; Sebanz, Knoblich & Prinz, 2003), and ascribing properties of ourselves to others (Otten & Epstude, 2006; Smith & Henry, 1996).

Consistent with a heuristic account (but importantly not inconsistent with a social attribution account), the truth bias is an objectively valid belief in the world outside of the controlled laboratory. Relational partners lie on average approximately two to three times per week (Guthrie & Kunkel, 2013). Diary studies suggest lying behaviour in everyday life is more frequent than this, occurring twice per day (DePaulo et al., 1996). One study found people lie less often to those who they felt closer (DePaulo & Kashy, 1998). Thus the base rates of lying behaviour outside of the laboratory may differ between those we consider psychologically closer to us and those we consider more distant. An ingroup bias may be adaptive where the base rates of honesty are naturally more frequent when interacting with ingroup rather than outgroup members. Thus expectations about others may be influencing the direction of judgment about others (see for example Allport, 1954; Asch, 1946; Fiske & Taylor, 1991; Hilton & von Hippel, 1996; Rosenthal & Jacobson, 1968; Snyder & Swann, 1978). Future research will need to determine what sorts of information is being inferred from mere grouping and how that comes to guide our lie-truth judgments.

One possible criticism that could be levied against the current studies is that individuals may anticipate reciprocity from ingroup members in future interactions, or perhaps fear outgroup discrimination of the ingroup (Moy & Ng, 2006): thus it was not mere relatedness that drove the findings but rather expectations about future interactions (Diehl, 1990; Tajfel et al., 1971). Whilst this may explain the presented findings, that the speakers in the videos were unidentifiable and that there was no obvious way in which the recorded speakers could reciprocate towards the participant or discriminate against ingroup members limits the utility of this explanation, although it cannot be ruled out.

Consider also raters believed they were grouped based on their artistic preference. They may have perceived similarities with ingroup members and dissimilarities with outgroup members based on artistic preferences and treated them differently on this basis (Allen & Wilder, 1979; Diehl, 1990). Note however prior studies have shown when the grouping procedure is random and that individuals are aware of the random arbitrary nature of the grouping, ingroup preferences remain (Billig & Tajfel, 1973). Thus it appears mere grouping is sufficient for an ingroup bias even when it is evident to participants that there is no basis for assuming similarity between ingroup members.

The bias toward believing ingroup members may be thought problematic for the group. Individuals sometimes choose to act in the interest of the self but to the detriment of their group by selfishly using up resources (Hardin, 1968). At the same time, they wish to remain part of the group whether for social or economic reasons. Free riders could exploit the truth bias of their ingroup peers without fear of group exclusion. The findings presented suggest raters can somewhat accurately classify behaviours of ingroup members, as indicated by the signal detection measures shown in Experiment 10. This suggests whilst individuals give their peers the benefit of the doubt, they are not naïve to the possibility of deception. This interplay between truth bias and classification accuracy might allow groups to strike a balance between maintaining social cohesion and reducing the risks of ingroup exploitation.

In conclusion, this chapter began asking how social information is used in the decision making process. Experiment 9 set out to test whether the socialisation practices we engage in mean we are more likely to judge others as truth-tellers, a claim that runs through the Ostrich effect (Vrij, 2008; Vrij, Granhag & Porter, 2010) and the accusatory reluctance account (DePaulo & Rosenthal, 1979; Ekman, 1992; Miller et al., 1986; O'Sullivan, 2003; O'Sullivan et al., 1988; Vrij, 2008; Vrij, Granhag & Porter, 2010). The findings suggest mere implied social presence is not sufficient to bring about a willing sacrifice of accuracy in favour of giving people benefit of the doubt. However, Experiment 10 shows that whilst mere presence is not sufficient, the nature of the relationship can mediate the presence or absence of the truth bias, independent of demeanour cues, speech cues, and so on. To what degree the ingroup truth bias is dependent on intragroup cohesive motivations, real-world base rates of honesty or social heuristics is yet to be elucidated, but these studies offer the initial steps towards exploring these accounts.

Chapter 7: General Discussion: Dual Process Theories

How do we make decisions in socially situated environments when there is little diagnostic information available to us? From research on language comprehension (Carpenter & Just, 1975; Clark & Chase, 1972; Clark & Clark, 1977; Gilbert, 1991; Mayo et al., 2004; Trabasso et al., 1971; Zwaan et al., 2002) to judgment and decision making (Bogacz, Brown, Moehlis, Holmes & Cohen, 2006; Bröder, 2003; Chaiken & Trope, 1999; Fox, 1957; Gilovich et al., 2002; Hall, 2002; Newman et al., 2012; Plott & Smith, 2008; Richter et al., 2009), it has been shown we make systematic decisions despite ambiguity in the environment. In a variety of research areas there is a bias towards accepting information to be true, known variously as the truthiness effect (Newman et al., 2012), the truthfulness bias (Zuckerman et al., 1979), the truthfulness heuristic (Burgoon, Blair & Strom, 2008), the truth effect (Koch & Forgas, 2012), and the truth bias (Bond & DePaulo, 2006). This thesis set out to explore the cognitive processes involved in decision-making in naturalistic uncertain environments.

Lie detection is an inherently social task, and one we have both an extensive history with (see DePaulo & Kashy, 1998; DePaulo et al., 1996; Park, Levine, McCornack, Morrison & Ferrara, 2002) and a wide range of beliefs about (see Hartwig & Bond, 2011; The Global Deception Team, 2006; Vrij, 2008; Vrij, Granhag & Porter, 2010). And yet despite our prior knowledge and beliefs we make for poor lie detectors (Bond & DePaulo, 2006; Kraut, 1980), at least in part because as liars we give off very few behavioural indicators of deceit, if any (Levine, 2010; Vrij, 2008;

Vrij, Granhag & Mann, 2010; Vrij, Granhag & Porter, 2010). For these reasons, lie detection offers a unique environment where raters bring with them prior knowledge and experience of the situation, along with an understanding of the social rules surrounding it, but nonetheless is a decision where uncertainty reigns. The mix of a complex set of information sources coupled with high uncertainty provides a relatively naturalistic environment in which to study decision-making under uncertainty. The core contribution of this thesis has been to explore precisely this.

In addition, lie detection research has rarely attempted to understand how the decision is formed and why raters so often go wrong (Miller & Stiff, 1993; Reinhard & Sporer, 2010; Vrij & Granhag, 2012). From this perspective, my thesis has contributed to the lie detection literature by showing how the current albeit limited view in favour of HAMs (Gilbert, 1991; Hawkins, Hoch & Meyers Levy, 2001; Masip et al., 2006; Masip, Garrido, et al., 2009; Masip et al., 2010; Reinhard & Sporer, 2008, 2010; Ulatowska, 2013; c.f. Roggeveen & Johar, 2002) is not warranted. I have introduced into this area of research a theoretical perspective borrowed from cognitive psychology: the adaptive decision-maker account (Payne et al., 1993; Simon, 1990). I believe this new perspective offers promise for future research into lie detection. From my research it appears raters make use of small amounts of readily available information in the environment; where such information is lacking, raters can draw on their prior knowledge of the world and of similar situations.

As well as these meta-theoretical contributions to decision-making and lie-detection research, my thesis offers four novel contributions to our understanding of socially situated decision-making under uncertainty. First, perhaps most importantly, I have shown that there is a cognitive component to the truth bias independent of the behaviour being rated, an assumption that has been made in prior research but not formally tested. In Experiment 1 I showed the initial truth bias and the subsequent decline in biased responding results at least in part from the rater independently of the behaviours of the speaker.

Second, in contrast to the current direction process-oriented lie detection is taking, I show a dual process model is *not* an adequate explanation of the phenomena. It failed to meet any of a set of challenges that I issued against them (Experiments 1, 2, 3 and 5). No support was found for the role of heuristic processing, but the use of heuristic rules of thumb built up from prior experience appear to be key to understanding how decision makers form their beliefs under uncertainty.

Third, the truth bias is characterised not as a flaw in an otherwise able decision-making system (Gilbert, 1991), but rather as a useful means of making informed decisions under uncertainty. I take an adaptive decision-maker position (Gigerenzer et al., 1999; Payne et al., 1993; Simon, 1990), claiming the truth bias is situation-specific and arises from an interaction between information available in the environment and that which we bring to the task. This is considered in detail in Chapter 8.

Finally, although the starting point of this thesis was to determine the cause of the truth bias, I conclude that the truth bias has no simple, single cause, but rather is an emergent property of the interaction between mind and world. That is to say that the truth bias is an outcome of this interaction, and that modifying either component can lead to a lie bias or no bias, demonstrating flexibility to the demands of the task and one's prior knowledge.

In this chapter I will summarise my research. First, a brief recount of the narrative will be given in the Summary of Research section. I will then discuss in more detail how my research contributes to our understanding of the truth bias. I will argue for a cognitive contribution to the truth bias that the rater brings to the task, independent of the behaviours being rated. Although there is a cognitive element to the truth bias, I will discuss how my research shows the HAM is not the causal factor.

Summary of Research

Rather than beginning from the assumption that the truth bias is a cognitive phenomenon, Experiment 1 (The Behavioural Account) asked whether the tendency towards believing others could be considered a bias or whether it was an accurate judgment of the apparently honest behaviours that truth-tellers exhibit and that liars attempt to display. It used the established phenomenon of the truth bias declining across successive ratings (Masip et al., 2006; Masip, Garrido, et al., 2009). It was shown that, independent of the behaviour being judged, there was an effect of the decision-making process that the rater brought to the task independent of the behaviours being rated. That is, regardless of the viewing order of behaviours, the reduced truth bias was observed across each new rating.

Experiment 1 was also tested the heuristic-analytic model (HAM), a dual-process theory that proposes a quick but error-prone heuristic processing route and a second more effortful but more deliberative analytical process (see Chapter 2 for a fuller description). While I have replicated the declining truth bias phenomenon across successive ratings, little support was gained for any of the class of HAMs identified by Evans (2007). The amount of viewing time was not able to predict the declining truth judgments, whereas the number of decisions made was (Experiments 1-3), in contrast to both default-interventionist and parallel-competition models.

Experiment 2 (Channel Effects) more directly examined the predictions of the pre-emptive competition model. This model predicts that the type of processing is either chosen early on due to particular preferences or is determined by the nature of the information available: behaviour requiring greater attention and deliberative interpretation such as speech cues require analytical processing (Chaiken, 1980; Evans, 2007; Evans et al., 1993; Forrest & Feldman, 2000; Gilbert & Krull, 1988; Reinhard & Sporer, 2008, Study 2; Stiff et al., 1989), whilst more stereotyped behaviours typical of visual behaviours (Forrest & Feldman, 2000; Reinhard & Sporer, 2008, Study 3; Stiff et al., 1989) are thought to be processed heuristically. However, the findings of Experiment 2 did not support this model, showing that regardless of the types of information available there was a decline in the PTJ.

Experiment 4 (The Consistency Heuristic) explored an alternative account of the decline. That the number of judgments made could predict the decline in the PTJ but that viewing time could not led to the consideration that raters could be making relatively simple *consistency* judgments between the responses presented. The findings supported this prediction: across the data collected in Experiments 1 (The Behavioural Account) and 3 (Thin-Slicing) as well as in previous research carried out by Masip et al. (2009), consistency was a reliable predictor of the decline in the PTJ, and indeed turned out to have diagnostic validity. Thus raters successfully made use of a relatively simple but informative cue. However, although raters made use of consistency, shown to be diagnostic, accuracy did not improve over time, possibly suggesting the integration of other cues in the decision-making process. Chapter 4 concluded that a dual-process theory was not supported by the data and suggested instead that in an uncertain world we may make use of few but reliable pieces of information as they become available.

The data so far do not entirely dismiss HAMs: the switch from heuristic to analytical processing may occur at a finer time scale than had been examined in my research to this point. Experiments 5-8 set out to test a prominent theory that has had an impact across a range of different research areas: the Spinozan mind hypothesis (Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993). It claims in order to be able to comprehend it is necessary to first accept the information as though it were true, at least in the first instance. Afterwards the rater can decide whether or not to continue believing the information, making this a dual process theory (Gilbert, 1991). This account of the truth bias assumes raters fail to update adequately from the initial automatic belief, which acts as an anchor (see Elaad, 2003; Fan et al., 1995; Zuckerman et al., 1987; Zuckerman, Koestner, et al., 1984).

Although the Spinozan account is consistent with prior findings (e.g. Carpenter & Just, 1975; Clark & Chase, 1972), I argued the data could equally be explained by a modified Cartesian account. This account claims there is no need to assign belief during comprehension: only after the information has been understood need a judgment of it be made. As such, it is a single process account (Gilbert, 1991). Provided we allow for the possibility that the Cartesian mind can use its experience with the world, the Spinozan effect of an early truth bias can be equally explained as a Cartesian guessing bias under uncertainty, making use of their prior knowledge to guide the judgment. By removing the necessity to respond, a Cartesian mind would no longer need to rely on its prior knowledge to make an informed guess and could instead simply indicate that it is unsure. Experiment 5 (The Cartesian Mind – Online Comprehension) found support for the modified Cartesian view, a single-process account, showing that removing the necessity for a judgment reduced the truth bias, particularly during the early moments of the decision process.

The prior knowledge relied upon may be base rate information: how often people usually tell lies and truths. People tell the truth far more often than lie (Caspi & Gorsky, 2006; DePaulo & Kashy, 1998; DePaulo et al., 1996; George & Robb, 2008; Hancock et al., 2004; see also Cole, 2001; van Swol et al., 2012), which could explain why raters are biased towards believing when forced into a judgment. Experiment 6 (Most People Tell the Truth – The Availability Heuristic) found that, when forced to judge early on, raters made use of base-rate information. This was taken as support for the contention that, when forced into judgment, raters rely on relevant situational knowledge. This study was the first investigation into the effects of perceived base rates on the truth bias.

The truth bias is found not only during the act of making a decision, but at the point of making a final lie-truth response (Bond & DePaulo, 2006; Zuckerman et al., 1979). Experiment 7 (The Cartesian Mind – Post Comprehension) sought to extend these findings to the point at which participants provided a response.

Surprisingly, and in contrast to what was predicted, raters showed a *greater* truth bias when they were not forced into judgment. Exploration of the data provided an explanation of this effect: raters were less likely to make lie judgments when able to indicate their uncertainty but showed no change in truth judgments, resulting in a greater proportion of truth compared to lie judgments. This artefactual truth bias may have been expected: prior research indicates raters are more uncertain when making lie than when making truth judgments (Anderson et al., 2002; DePaulo et al., 1997; Anderson, 1999, cited by DePaulo & Morris, 2004; see also Levine et al., 1999), and that raters are more unsure when they are actually listening to a lie than when listening to a truth (Anderson et al., 2002; DePaulo et al., 1997; Hurd & Noller, 1988). That is, internal uncertainty is diagnostic of deception.

Although the following explanation results from an exploration of the data, and so caution must be taken, it is nonetheless promising to note that, when forced into making a judgment after information from the environment has been obtained, raters can make use of their uncertainty in an adaptive way. If the available information fails to lead to a truth judgment, the behaviour is likely indicative of a lie. Given that truths are easier to detect, this situation-specific knowledge can be used to suggest a statement is deceptive when confidence is low. This adaptive use of uncertainty can be seen in the comparison between the forced and unforced choice condition: raters who were forced to judge were more likely to judge statements to be deceptive but no more or less likely to judge statements to be truthful. Adaptive decision-making is the theme of Chapter 8 and is discussed further there.

Exploring the data of Experiment 7 showed how uncertainty can be used as a diagnostic indicator of deception: when raters were uncertain but forced to make a decision, they opted to make a lie judgment. But Experiment 5 found what seems at first to be a contradictory result: when raters were unsure, they were more likely to make a truth judgment. I claimed the reason for these differences rests on a distinction between what I have called internal versus external uncertainty.

Internal uncertainty results when it is difficult to make a decision *in light of* the evidence. The uncertainty arises because the evidence does not allow the rater to make a clear judgment, and so in this sense the source of uncertainty can be attributed internally to the rater's inability. Because lies are more difficult to detect than truths (Levine et al., 1999), internal uncertainty is more likely when listening to a lie, and can be used to make an informed guess that the statement is a lie.

External uncertainty arises from a lack of information in the environment. It is not due to indecision on part of the rater, but rather the sheer absence of any information on which to make a judgment. It is not a difficulty in weighing up the evidence, but rather simply a lack of evidence to be weighed. In short, internal uncertainty can be thought of as 'I cannot decide with this information', whereas external uncertainty can be thought of as 'I do not have enough information to decide'.

I argue that where uncertainty results from a lack of information in the environment (i.e. external uncertainty), raters rely on prior knowledge of similar situations. If a decision is to be made in the absence of information, raters do not simply guess randomly but try to make an informed guess from prior knowledge, such as base rate knowledge. This, it is suggested, leads to a truth bias, although if base rate knowledge suggested people usually lie, then a lie bias would be observed. In fact, this is what I demonstrated in Experiment 6.

However, when the information has been gathered about the statement but that information still leaves the decision-maker uncertain, this internal uncertainty can be used in an adaptive way to suggest the speaker is deceptive. That is because uncertainty at the point of the final judgment turns out to be diagnostic of deception: we tend to be more uncertain when listening to lies (Anderson et al., 2002; DePaulo, 1992; DePaulo et al., 1997; DePaulo et al., 2003; Hartwig & Bond, 2011; Hurd & Noller, 1988).

Common to both internal and external uncertainty is that situation-specific knowledge, whether in the form of base-rates or meta-cognitive awareness of the difficulty in spotting lies versus truths, can be used to guide the final judgment under uncertainty. The distinction between them is important because it shows that raters make use of different strategies dependent on the information available to them. Such flexible decision-making strategies illustrate the adaptive nature of making lie-truth judgments, a claim I pursue further in Chapter 8.

To carry out a confirmatory test of the distinction between internal and external uncertainty, and to more directly test the claims of the Spinozan mind account, Experiment 8 (The Cartesian Hopi Word Experiment) replicated Gilbert et al.'s (1990) first experiment with an added manipulation: participants in one condition were forced into a judgment (as per the original study) while those in the other condition were not. An interesting prediction arises from the adaptive decision-maker account. Whereas the forced choice condition should replicate the original findings of Gilbert et al. (1990, Study 1: a truth bias when judging previously interrupted statements and a reduced biased for uninterrupted statements), the unforced choice condition should show a reversal of the effect: a reduced bias when judging previously interrupted statements (external uncertainty cannot be used as a diagnostic indicator of deception) whereas uninterrupted statements ought to be truth biased (internal uncertainty can be used as a diagnostic indicator of the falsity of the statement because the available information failed to result in believing the statement). The Spinozan bias was replicated. However, there was no crossover effect: participants not forced into making a choice were truth biased whether they had been interrupted or not during learning.

On further examination, it became clear the task was not well situated for testing the Spinozan mind or the adaptive decision-maker accounts. The Spinozan effect could be better explained as a tendency towards forgetting interrupted false statements. Others have also noted the differences in ease of encoding, showing the encoding of false statements ('a twyrin is not a doctor') is difficult because the various possible interpretations of the true definition of the word (a twyrin may be anything other than a doctor) need to be processed in a limited time period (Anderson et al., 2009, 2010; Glenberg et al., 1999). As it stands, this thesis offers a novel account of the truth bias resulting from an interaction between the information in the environment and one's knowledge of similar contexts.

Experiment 9 (Socialisation Practices) considered whether the artefactual truth bias found at the point of the final judgment, as in Experiment 7, results not from an adaptive use of internal uncertainty but rather from social practices. It is socially aggressive to confront someone and to call them a liar, and may result in an equally aggressive response (DePaulo & Rosenthal, 1979; Ekman, 1992; Miller et al., 1986; O'Sullivan, 2003; O'Sullivan et al., 1988; Vrij, 2008; Vrij, Granhag & Porter, 2010). In addition, there are social benefits to choosing not to confront a lie: challenging the claims of others can cause relationships to break down (Bell & DePaulo, 1996; Clark & Lemay, 2010; Cole, 2001; DePaulo & Bell, 1996; Guthrie & Kunkel, 2013; Miller et al., 1986), and often people lie to us out of politeness, wanting not to hurt our feelings (DePaulo et al., 1996; Lewis, 1993; Vrij, 2008; Vrij, Granhag & Porter, 2010), and so we may choose to actively avoid spotting lies (Vrij, 2008; see also DePaulo et al., 2003; Ekman, 1992; Vrij et al., 2010).

I examined whether there was a social basis to the truth bias. Accounts of this nature claim at their core that mere implied social presence is sufficient to result in a truth bias, and in turn a sacrifice of accuracy. I found no support for this claim: regardless of whether the speaker was believed to be a social or non-social agent, there was no truth bias. It is not possible to determine the cause of the lack of truth bias. Minimal information was presented with the intention of discovering the necessary conditions for the truth bias. Thus the lack of bias may reflect the lack of video or audio behaviours, a lack of familiarity with the stimuli presented, and so on. For the current purpose it is important to note that the implied social presence of another was not sufficient to invoke a truth bias, as has been suggested previously.

Previous research has shown how social information influences truth responding: we are more inclined to believe those to who we feel close rather than strangers (Argyle & Henderson, 1984; McCornack & Levine, 1990; McCornack & Parks, 1986; Stiff et al., 1992; Stiff et al., 1989; van Swol et al., 2012; see also Boyes & Fletcher, 2007; Gaertner & Dovidio, 2000; Voci, 2006). It was suggested this situation-relevant social context information could be used in order to aid the decision-maker. It is the case that we tend to tell the truth more often to ingroup compared to outgroup members (Cole, 2001; DePaulo & Kashy, 1998)[7], so this would prove to be a useful, adaptive strategy. A minimal grouping paradigm was employed in Experiment 10 (Social Relatedness as a Heuristic) to examine the effects of relatedness. A minimal paradigm was used because previous studies have confounded social relatedness with familiarity with the speaker. A truth bias was observed when rating ingroup members, but there was no evidence of a bias when judging outgroup members.

In summary, little support was found for any of the three classes of HAMs identified by Evans (2007). Instead, this thesis suggests that in forming beliefs we make use of a relatively diagnostic heuristics in the form of prior knowledge, an adaptive use of internal uncertainty, consistency information or social information. The use of these heuristics was seen to be dependent on the structure of the task. I

---

[7] Although the types of lies that people tell to close relations may be more serious in nature than those they tell to acquaintances or strangers (L. J. Speed, personal communication, August 12, 2013; DePaulo, Ansfield, Kirkendol & Boden, 2004)

adopt Simon's (1990) view of rationality in this thesis as a scissors with two blades, one representing the structure of the environment and the other our limited computational ability. Belief formation cannot be understood as either one or the other: both blades of the scissor are necessary, discussed further in Chapter 8.

Belief Formation as a Dual Process

Having given a recount of the findings of my research, I will now focus on whether they evidence a single or dual process account. To do so, I will first consider whether my research can be reinterpreted as showing that the truth bias is actually nothing more than a reflection of the apparently truthful behaviours the speakers give off. I will argue that it is not possible to ignore the independent cognitive component of the truth bias that the rater brings to the task. Having established this, I will move on to integrate my studies with prior research to arrive at the conclusion that a dual-process model fails to meet empirical challenges and is conceptually flawed.

The Truth Bias: A Cognitive Phenomenon

Experiment 1 sought to determine whether the truth bias could be explained simply as a reflection of the behaviours that are displayed by truth-tellers and that liars attempt to convey in their self-presentation. It was shown that regardless of whether a speaker's initial response or final response was rated first, there was a truth bias. Across multiple judgments the bias was seen to decline. This was taken as support for an independent cognitive component to the truth bias.

Yet Experiment 4 found that consistency in the speaker's response, a diagnostic indicator of their dishonesty, was able to account for the decline in the bias over time. Thus the phenomenon could be explained with recourse to the consistent or inconsistent behaviours displayed across the course of their statement. Consider also that Experiment 2 found an increase in accuracy with longer first responses of the speaker when given only audio information. One possible account of the finding suggested in the discussion of Experiment 2 was that the serial nature of speech, necessarily a temporally-extended cue, may account for the increasing accuracy: over time, more information became available to the rater. Similarly, in Experiment 3 accuracy increased for short clips between viewing 8 and 16 s, suggested to be due to the short amount of information available after 8 s. Again, there appears to be an exogenous influence on the judgment outcome.

An important question is whether there is any evidence of an endogenous role in accounting for the truth bias. Experiments 5 and 7 provided evidence that, despite judging the same behaviour, the degree of belief in the statement was dependent on whether raters were able to explicitly indicate their uncertainty. This was particularly true during the earliest moments of comprehension in Experiment 5. Perhaps most convincingly in favour of a cognitive component to the truth bias, Experiment 6 found that beliefs about the likelihood with which a person was lying resulted in raters showing a truth or lie bias as the statement was being delivered. Similarly, Experiment 10 found that mere grouping of speakers as in- or out-group members determined whether raters showed a truth bias or not. Thus there is clear evidence in favour of endogenous forces on the truth bias. What form these forces take is considered next.

Dual or Single Process?

Process-oriented lie detection research has begun to take a dual-process approach (e.g., Masip et al., 2006; Masip, Garrido, et al., 2009; Reinhard & Sporer, 2008, 2010). In one line of research, belief formation is seen as a default-interventionist system that initially (and automatically) believes but is later interrupted by a more effortful evaluation (Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993; Masip et al., 2006; Masip, Garrido, et al., 2009; Masip et al., 2010). A second pre-emptive approach is considered later. Given the prominence of default-interventionist models, this thesis carried out a number of studies examining their predictions, but found them wanting (Experiments 1-3, 5, 6-8).

This is not the first time the default-interventionist HAM has been challenged. Deutsch et al. (2006) note that with sufficient practice, the processing of negated statements, expected to take longer because of a necessity to update from automatic believing, can be as quick as processing affirmed statements. Although this can be incorporated into dual-process theories, assuming that more effortful processing can become sufficiently learnt to be automatized and make use of a more heuristic mode of thought (Deutsch et al., 2006), it is difficult to align with the stronger position that proposes disbelieving *necessarily requires* an additional tagging after the mental construction of the statement as true (Gilbert, 1991). More difficult for the more general class of default-interventionist models, mere priming of a negation (rather than long-term training) is sufficient to result in unintentional (and thereby automatic: Bargh, 1989), fast and effortless processing of negated statements (Deutsch et al., 2009; Dodd & Bradshaw, 1980; Schul et al., 2004). Rigorous practice is not required for apparent automaticity in the processing of negated statements, suggesting analytical processes can be as fast as heuristic processes.

If negated statements can be processed as quickly as affirmed statements, why has research so often found an affirmed-statement advantage (Carpenter & Just, 1975; Clark & Chase, 1972; Clark & Clark, 1977; Donders, 1969; Gilbert et al., 1990; Mayo et al., 2004; Trabasso et al., 1971; Zwaan et al., 2002)? It has been suggested that negated sentences take longer to process because they are typically underspecified, and so it is a quirk of the task structure that gives the misleading impression of a delay in processing false statements (Anderson et al., 2010; Glenberg & Robertson, 1999). That 'the eagle is not in the sky' leaves open a large array of possible places that the eagle may be: in its nest, on the ground, and so on. 'The eagle is in the sky', the affirmative, rules out these alternatives. Theories of language comprehension have proposed understanding entails an online mental simulation of the state of affairs depicted in the statement (Glenberg et al., 1999; Zwaan & Radvansky, 1998; Zwaan et al., 2002). As such, the underspecified (negated) statements are liable to take longer in order to process the various alternatives, i.e. with the eagle on the ground, in its nest, and so on (Anderson et al., 2009, 2010; Glenberg et al., 1999).

This fits well with the view that comprehension, necessarily an online task, requires a constant evaluation and verification of the sentence details (Richter et al., 2009; Schroeder et al., 2008), which Richter et al. (2009) have called *epistemic monitoring*. Consider for example the processing of the locally ambiguous sentence "The detective called the suspect the officer caught a liar". If difficulty is incurred part way through the sentence, readers can re-evaluate by scanning backwards through the message (Altmann, Garnham & Dennis, 1992; Frazier & Rayner, 1982; Rayner, 1998; although see Mitchell, Shen, Green & Hodgson, 2008). Comprehension appears to be effortful and corrective. There is also a body of research demonstrating the online effects of competition when processing sentences and even words with multiple meanings (see Harley, 2008; Spivey, 2007; Spivey, Tanenhaus, Eberhard & Sedivy, 2002), again suggesting that comprehension is an online *evaluative* process, not an automatic acceptance of all information. Evaluation is not a property of the faster heuristic processing stream, but rather of the slower analytical stream. Thus the evidence suggests a routine use of analytical processing that operates extremely quickly, at the level of phonemes (Spivey et al., 2005), in order to constantly monitor the comprehension process. This, along with other downfalls of the strong default-interventionist position of the Spinozan mind (Gilbert, 1991), is considered in more detail in the next section.

A second dual-process approach to understanding lie-truth judgments has applied pre-emptive conflict resolution models, albeit with a focus on accuracy rather than judgmental bias (Reinhard, 2010; Reinhard & Sporer, 2008, 2010). The findings of this thesis did not show support for a role of this class of models in explaining the truth bias. The amount of processing time available should be a useful guide as whether to select the faster heuristic processing route or the slower analytical route, but this was not supported (Experiment 3). More conclusively, when the environment is structured such that analytical or heuristic processing is necessarily *required*, the selection of the relevant processing stream should be made from the outset, but again this was not found (Experiment 2). These findings are consistent with research showing that people do not give more time to processing conclusions that are relatively difficult to believe (Ball et al., 2006; Evans, 2007; Thompson et al., 2003), as would be predicted by pre-emptive conflict resolution models: again, believability should be a readily accessible cue as to whether to select heuristic or analytical processing (Evans, 2007).

So far the evidence does not appear to support particular instantiations of the dual-process heuristic-analytic model. Dual process models more generally have also come under criticism. A number of studies that have shown how different information types, assumed to engage heuristic or analytical processing, have been confounded with task demands: message content cues, thought to encourage analytical processing, are also lengthier, more complex, and are delivered later in the information stream compared to more readily accessible and shorter visual cues (Chun & Kruglanski, 2006; Erb et al., 2003; Kruglanski et al., 2006; Kruglanski & Thompson, 1999a, 1999b; Pierro, Mannetti, Erb, Spiegel & Kruglanski, 2005). The difference in judgment outcomes when making use of these cues may reflect shifts between heuristic and analytical processing, but could equally be accounted for as a change in the degree of cognitive load placed on the individual. The judgmental outcome differences disappear when these confounds are removed (for a review, see Kruglanski et al., 2006), suggesting cognitive load is an important predictor of the final judgment, more so than the distinction between heuristic and analytical processing.

These empirical findings are difficult to explain under a dual process account. Dual process accounts also encounter difficulties on a conceptual level. Kruglanski et al. (2007; see also Gawronski et al., in press) note the various conceptualisations of heuristic processing made by different theories (e.g., Chaiken et al., 1989; Petty & Cacioppo, 1979; Smith & DeCoster, 2000; Strack & Deutsch, 2004) do not readily map onto one another because they make quite different claims as to what and how information is processed by the heuristic or analytical processing stream (Gigerenzer & Regier, 1996; Newstead, 2000; Osman, 2004; Stanovich & West, 2000). If each of these alternative views of the heuristic process are supported by research, there must be a vast array of processing methods available to us, from the ‘peripheral’ (Petty & Cacioppo, 1986) and ‘associative’ (Smith & DeCoster, 2000), to the ‘impulsive’ (Strack & Deutsch, 2004) and ‘heuristic’ (Chaiken et al., 1989), see Evans (2008).

It is not only difficult to find the overlap in the different conceptualisations of heuristic processing *across* dual-process models, but also *within* any given model the boundaries between heuristic and analytical processing are difficult to find. For example, heuristic processes are said to be effortless and fast while analytical processes are effortful and slow. These two variables, of effort and speed, are continuous in nature. It is not clear where the line should be drawn to separate effortless from effortful processing, fast from slow. A distinction of two separate processing modes must show non-overlap between the processes if they are to offer an explanatory advantage over single process models (Keren, 2013; Keren & Schul, 2009; Osman, 2004), but this is both difficult to test empirically (Keren & Schul, 2009; Schacter & Tulving, 1994) and leads to the unlikely logical outcome (according to Keren & Schul, 2009) that each processing route can form a complete judgment in isolation of the operations that are solely within the remit of the other processing route (Keren, 2013; Keren & Schul, 2009; although see Evans & Stanovich, 2013). That is, the two modes of processing, if they are to be considered conceptually distinct, must be shown to be functionally distinct and mutually independent.

However, Kruglanski and colleagues (2006; 2007) review research showing how phenomena classically considered as resulting from heuristic processing has now been recognised to be evaluative, logical rule-based, and therefore analytical. They also show that some analytical rule-based decision-making can become automatized, a property of heuristic processing (see also Deutsch et al., 2006). The overlap between the two processes may be thought of as an unfortunate and somewhat unsurprising result of dichotomising a continuous variable. Consider, for example, that the heuristic processing stream has been considered the fast and effortless process and analytical processing the slow and effortful process. These variables, of speed and effort, are evidently continuous in nature. Whether it makes sense to attempt to draw a line at a particular speed, at which everything below is considered slow and everything above is considered fast, is at best debatable (Osman, 2004). The findings of a number of reviews argue this position (Gawronski et al., in press; Keren & Schul, 2009; Kruglanski et al., 2006; Kruglanski et al., 2007) and suggest that a binary distinction between two modes of processing is not a valid one.

The findings reviewed above along with those presented in the current thesis suggest an alternative approach is needed. The data better align with a single process model, such as the unimodel of decision making (Kruglanski et al., 2006; Kruglanski et al., 2007). The unimodel claims decisions result from a single processing stream that is part determined by both the (internal) available cognitive resources and (external) task demands, echoing the bounded rationality claim of an interaction between the mind and the environment (Chase et al., 1998; Payne et al., 1993; Simon, 1990; Todd & Gigerenzer, 2005).

That different information types (i.e., external task constraints) or degrees of cognitive load (i.e., internal cognitive resources) influence the trajectory of the final decision is thought not to reflect distinct and separate cognitive structures (as per a dual process account), but rather to how parameters within the model are tuned depending on context and the availability of cognitive resources. A number of belief formation models have taken the single-process view and have similarly conceptualised the process as one that varies along a continuum, with parameters that can be tuned to the task (Fiske & Neuberg, 1990; Kahneman, 2003; Kruglanski et al., 2007; Petty & Cacioppo, 1986; Sherman, 2006). For instance, cognitive load can make it more difficult to remember the relevant information (see Experiment 8), which in turn leads to different judgment outcomes despite arising from the same processing system. The adaptive decision-maker account takes a similar perspective, arguing that decision outcomes are the result of the interaction between situational context and the limited internal cognitive resources (Experiments 5-7).

Although a single process account can also explain the findings, should a single process model be *favoured* over dual process models when it comes to belief formation? The current thesis argues so. Phenomena purportedly supporting a dual process position could not only be equally explained by a single process model making use of adaptive rules (Experiments 4-6), but it was also shown that an adaptive decision-making account is able to avoid the conceptual difficulties associated with dichotomising continuous variables and is able to make novel predictions based on the context-sensitivity of the decision process. Partial support was found for the context-dependency of decision-making under uncertainty (in particular, see Experiments 5 through 8). Full support could not be attained because of the unexpected flaw in the design of the extended replication of Gilbert et al. (1990, Study 1), a flaw originating from the initial experiment, namely that participants were more likely to forget the less informative false items under time pressure (see Hasson et al., 2005; Kruglanski et al., 2007). Nonetheless, those predictions that were made were substantiated (Experiments 4, 5, 6, 9 and 10): simple context-dependent rules are used to guide decisions under uncertainty. There was one apparent exception: Experiment 7 produced a result that was in the opposite direction of the prediction. Raters became *more* truth biased if they were not forced into judgment. In light of prior research, such an effect may have been anticipated. Given that raters are more confident in making truth judgments and are more confident when rating truths (Anderson et al., 2002; DePaulo, 1992; DePaulo et al., 1997; DePaulo et al., 2003; Anderson, 1999, cited by DePaulo & Morris, 2004; Hartwig & Bond, 2011; see also Levine et al., 1999), it is perhaps not surprising that, if forced into judgment, uncertainty itself can be taken as an indication of deception. However, this is a post-hoc account of the findings and requires a confirmatory test in future research. Nonetheless, a single process account of belief formation offers substantial promise.

Although there has been no support shown here for a dual process theory of the truth bias, it may be argued that it offers researchers a means of making and testing predictions. As discussed, recent research has found the pre-emptive conflict resolution model a fruitful theory for making predictions about lie detection accuracy (Reinhard, 2010; Reinhard & Sporer, 2008, 2010), although it should be noted that Reinhard and Sporer (2008) state their findings are not in contradiction to a single-process models. Thus the evidence seems to strongly point towards the absence of a dichotomising of mental processes and towards a single mental system.

For now, this thesis takes the single-process position of Kruglanski and colleagues (2007, p. 296):

> '[It is the] parametric differences between the… degree of *task demands* or of *subjective relevance of information*… rather than other possible distinctions (e.g., in the type of contents of the information processed, awareness, or swiftness of processing) that account for the empirical results on which numerous dual-mode formulations were based.' (Emphasis added).

Reinterpreting Evidence Supporting a Dual Process Position: The Spinozan Mind

I have argued that HAMs fail to meet empirical tests, as shown in my research, as well as suffering from conceptual flaws that are the source of on-going debate (Evans & Stanovich, 2013; Keren, 2013; Keren & Schul, 2009; Kruglanski et al., 2007; Osman, 2004; Thompson, 2013). One default-interventionist HAM has been particularly prominent in lie detection research: the Spinozan mind. In contrast to the above discussion, it has been claimed that the Spinozan mind is supported from a variety of different research areas and exhibited in a range of phenomena (Gilbert, 1991). In this section I will reconsider the evidence that has been shown to support the Spinozan mind account and illustrate how it may be accounted for under a single processing view.

The main form of evidence in support of a Spinozan position is that comprehension of affirmed statements (e.g. 'star [symbol] is above plus [symbol]') is faster than comprehension of negated statements (e.g. 'star is not above plus': Clark & Chase, 1972; see also Donders, 1969; Zwaan et al., 2002). In Chapter 2 and in the above section I showed how this evidence is compromised by the fact that negated statements in these studies have a greater number of possible interpretations: e.g. the star may be below the plus, behind it, and so on (Anderson et al., 2009, 2010; Glenberg et al., 1999; Hasson et al., 2005). That is, this finding was the result of the structure of the task environment: when negated statements have a similar number of possible interpretations (e.g. 'this coin does/does not show heads to be face up'), participants were equally fast to respond to negated and affirmed statements (Anderson et al., 2009, 2010; Glenberg et al., 1999). This evidence has been used to argue for a single-process model that is slower in processing statements that require greater consideration (Anderson et al., 2009, 2010; Glenberg et al., 1999).

A second stream of research, the confirmation bias (covered in Chapter 2), has also been taken as support of the Spinozan position (Gilbert, 1991) insofar as people show a bias towards seeking evidence in favour of believing over disbelieving. Despite being heavily documented (see Snyder & Campbell, 1980; Snyder & Swann, 1978), evidence in favour of an ever-present confirmation bias is lacking (Higgins & Bargh, 1987; see also Trope & Bassok, 1982). In a review of the area, Higgins and Bargh (1987) concluded that participants only seek to confirm their hypotheses in situations where confirmatory questions can generate more information than can disconfirmatory questions (see Chapter 2 for a more detailed recount). Again, it seems we make use of contextual information in order to best guide our judgments.

Having established belief in a statement, it continues to persevere despite being told the statement was false (Nisbett & Ross, 1980). This has been considered a potential cause of the truth bias and as support for the Spinozan mind hypothesis, showing how early automatic assent of a statement continues to influence the judgment at later stages (see Elaad, 2003; Fan et al., 1995; Gilbert, 1991; Zuckerman, Koestner, et al., 1984). But belief perseverance does not reflect some miserliness or anchoring to an initial default position. When the opportunity for additional cognitive work is given, the initial belief is found to be *stronger*, not weaker (Fleming & Arrowood, 1979; Hovland, Lumsdaine & Sheffield, 1949; Skurnik et al., 2005), suggesting that belief perseverance results from further evaluative processing, not from a lack of exertion to overcome an initial automatic belief.

Perhaps most difficult to account for is the research Gilbert and colleagues (1990) muster to show that, despite being *forewarned* that they were about to hear a false statement, participants continue to automatically believe statements early on. This would seem to be strong support for the automaticity of believing. Although the authors empirically confirm the claim (Gilbert et al., 1990, Study 2), the prior research does not lend itself to this interpretation. Prior research (Allyn & Festinger, 1961; Hovland & Weiss, 1951; Kiesler & Kiesler, 1964; McGuire, 1964) and at least two meta-analyses (Benoit, 1998; Wood & Quinn, 2003) conclude that forewarning can in fact induce what Gilbert and colleagues (1990; Gilbert, 1991) refer to as a sceptic's set, showing that forewarnings do in fact make it easier to discredit statements as false. When forewarned, there does not appear to be the same tendency towards believing information to be true, contradicting the Spinozan mind claims.

Notably, potential moderators of the forewarning effect have been suggested: whether the warning was of an impending persuasion attempt or a more detailed warning about the contents of the message (Cialdini & Petty, 1981; see also Baron, Baron & Miller, 1973), and whether there are sufficient cognitive resources available (Kiesler & Kiesler, 1964; Wood & Quinn, 2003) affects the efficacy of the warning. Although these findings suggest a context- and cognitive resource-dependency, in line with an adaptive decision making account, that cognitive load influences the degree of believing could be taken to support the Spinozan position. According to this theory, if raters are not given sufficient time or have insufficient resources to move to the analytic evaluation of the comprehended statement, they will continue to believe. Gilbert and colleagues (1990, 1993) also show that cognitive load increases the propensity to believe a statement is true. That cognitive load increases the likelihood of believing was an argument that Gilbert (1991) used in support of the Spinozan position, because automatic processes (that cause an initial belief in the statement) should not be affected by cognitive load but more effortful processes (required to update the initial automatic belief) should be, resulting in a greater influence of the initial automatic belief stage.

As discussed in the preceding section, however, this cognitive load prediction is not unique to the dual-process Spinozan account. The single-process model of Kruglanski et al. (2007) also predicts task difficulty will influence the outcome of the decision process by influencing the degree to which relevant information can be encoded. As shown in Experiment 8 (see also Hasson et al., 2005), cognitive load interfered with the encoding of those statements that were under-specified, i.e. false statements. As such, the truth bias could be explained not as a tendency towards automatically believing, but rather as a difficulty in encoding those statements that required greater processing under time pressure. That cognitive load can result in a truth bias, resulting from a greater tendency to forget false items, both discredits the Spinozan view and can account for the findings presented by Gilbert and colleagues (1990, 1993) showing that cognitive load results in a bias.

It would seem from the my discussion that evidence employed as support for a dual-process Spinozan hypothesis can also be explained under a single-process Cartesian account. Indeed, as the original authors note, ‘most of these results [from prior research] can be explained within the Cartesian as well as the Spinozan framework’ (Gilbert et al., 1993, p. 222). But more than this, the adaptive decision-making account can also account for context-dependent and resource-dependent moderators, considered further in the next chapter.

Chapter 8: General Discussion: Adaptive Decision Making

I have argued that dual process theories cannot account for the truth bias and that they suffer from conceptual difficulties. Consideration is now given to whether the truth bias can be better thought of as an adaptive, successful strategy in an information-limited world, rather than as an unfortunate erroneous consequence of forming the difficult lie-truth judgment. I argue the presence of a truth bias is dependent on the types of information available in the environment: it is a flexible response to the types of cues available.

Adaptive Decision Making: Interplay Between World and Mind

This thesis opened with a distinction between heuristics and heuristic processing, which have previously been used interchangeably (Gigerenzer & Gaissmaier, 2011). However, I have shown that whereas people use relatively simple *heuristic rules* in making lie-truth judgments (Experiments 4, 6, 7 and 10), *heuristic processing* does not seem to be in play (Experiments 1-3). This distinction between simple heuristic rules and a heuristic processing mode is therefore an important one.

The definition of the truth bias I have used is an operational one of rating statements as truths more often than they are present. Some have taken a different definition, arguing the truth bias is a heuristic rule (Fiedler & Walka, 1993; Stiff et al., 1989), reflecting a generalised tendency to believe others as the name implies. The findings of this thesis suggest both an operational and a heuristic definition are in part true. The truth bias, I argue, can be thought of as an emergent property of an adaptive system. In making simple but effective decisions in an information-limited world, where liars provide very little indication as to their true intentions, I suggest we make use of readily available information in whatever form that may take, whether it is our prior experience with the world in the form of base rates (Experiment 6 of this thesis; O'Sullivan et al., 1988), social information in the sense of psychological distance (Experiment 10 of this thesis; Vrij, Granhag & Porter, 2010), our own uncertainty as resulting from an inability to decide (Experiments 7-8), or other salient cues available in the environment (Platzer & Bröder, 2012; Street & Richardson, in prep). Across the experiments reported herein, it was shown how such relevant contextual information could be used to guide the decision under uncertainty. That is, the source of the truth bias was dependent on the types of information available in the environment.

Two important considerations arise from the adaptive decision making account offered. First, if the truth bias results from an adaptive use of contextual and environmental information, this suggests raters will not always demonstrate a truth bias. Rather, flexibility in the types of information used could also lead to a lie bias. Do raters show such flexibility?

Second, if raters make satisfactory judgments using limited cognitive resources in a limited but information-rich world (Simon, 1990), do they employ compensatory or non-compensatory strategies? The former results when the additive effects of the cues used in the decision result in a net effect where cues can cancel each other. That is, a host of weaker cues suggesting deception could override the power of a single stronger cue that suggests honesty. A non-compensatory strategy describes the situation where cues do not have a net effect, but rather the most informative single cue is used in the decision and all other information has no effect, such that a set of weaker cues could not come to overpower the effect of the selected cue. These two questions are considered below.

Flexibility in the Truth Bias

With regards to the first concern, the evidence suggests people do not always show a bias towards believing others. Meissner and Kassin (2002) documented an investigator bias in police interviewers, who show a reliable tendency towards ratings speakers as liars (see also Bond et al., 2005; Kassin, 2005; Kassin & Fong, 1999; Kassin et al., 2005; Masip et al., 2005). One may suspect their line of work requires a degree of suspicion. Indeed they report that they expect their interviewees to lie to them when asked in research interviews (Moston et al., 1992), so this cannot be said to be evidence of flexibility in their responding. Rather, they overgeneralise their suspicion (see Kassin et al., 2005; Masip et al., 2005). But even ‘naïve’ undergraduates, who typically show a truth bias (Bond & DePaulo, 2006; Levine et al., 1999; Zuckerman et al., 1979), can in certain circumstances exhibit a lie bias. In the current thesis, it was shown how base-rate information is used early on in the judgment forming process (Experiment 6). When the base rates suggested 80% of the speakers were likely to lie, I found an initial bias towards disbelieving. By the point of their final judgment all respondents, regardless of their beliefs about the base rate, shifted more towards believing the speaker. That is, those who began with a lie bias at the start of the statement finished with no response bias by the end of the statement. Even across the course of a single statement raters show flexibility in their responding. Undergraduates given bogus training on the cues to honesty versus the cues to deception show a bias towards believing and disbelieving, respectively (Masip, Alonso, et al., 2009; see also Blair, 2006, for similar findings with police officers). Making raters more aware of cues to honesty or deception may have made them more salient than other behaviours and resultantly had a greater impact on the decision process (Masip, Alonso, et al., 2009; Platzer & Bröder, 2012). In two studies not reported in this thesis, we similarly found raters who attended to a cue known to suggest deception or honesty shifted the degree to which raters were prepared to believe the speaker (Street & Richardson, in prep).

Finally, by leading participants acting as interviewers to expect their partner to be guilty, they choose more guilt-presumptive questions: even third-party independent raters considered the speaker to be guilty (Hill, Memon & McGeorge, 2008), showing a self-fulfilling prophecy (Merton, 1948). Thus our prior beliefs can come to influence how we interpret behaviour. There is no question that lay people are typically truth biased. But there are situations where it is adaptive to have a preference towards disbelieving others: whether that is because we have reason to suspect others or because the most salient cues in the environment are those that suggest deceit.

It would appear then that the truth bias does not have a single cause, but rather emerges from an interaction between context-specific knowledge and information in the environment. I argue that the truth bias is an adaptive view of the world, reflecting the base-rate of honesty that laypeople typically encounter (i.e. that people tend to tell the truth: DePaulo & Bell, 1996; DePaulo & Kashy, 1998). When situation-specific knowledge or salient cues in the environment suggest otherwise, raters can adopt a preference towards disbelieving.

Compensatory Decision-Making

It is unclear whether the adaptive decision-maker incorporates a set of cues into the decision-process, which can have additive or subtractive effects on the judgment outcome, or whether a single cue is selected from amongst the set of available cues and is the sole basis for judgment, such that other available cues do not have an additive or subtractive effect. The former, a compensatory strategy, has the benefit of being more informed insofar as it makes use of more of the available information. The latter, a non-compensatory strategy, has the benefit of simplifying the complex decision-making process. The findings of the current thesis do not readily side with one account or the other. However, I have advocated an adaptive decision-making perspective in understanding decision-making in uncertain but information-rich worlds. Brief discussion is given to the matter here because it is an important consideration for any adaptive decision-making account including my own (see Bröder & Eichler, 2006; Bröder & Schiffer, 2003; Newell & Fernandez, 2006; Pohl, 2006), and is a current source of debate (Gigerenzer & Brighton, 2009; Hilbig & Richter, 2011; Newell & Shanks, 2003; Newell, Weston & Shanks, 2003; Oppenheimer, 2003; Pohl, 2011). For a review of the debate, see Pohl (2011).

A number of studies have shown that, where non-compensatory strategies are used (Gigerenzer, Hoffrage & Kleinbölting, 1991), they are used when participants are not free to search for information in their own way (Bröder, 2003; Newell & Shanks, 2003), and while the majority seem to use the rule a large proportion of people do not (Bröder, 2000; Newell & Shanks, 2003; Newell et al., 2003). For instance, it might be expected that the recognition heuristic would be used as a single noncompensatory cue: ‘choose the recognised item when unsure’. But some research shows it is used in conjunction with other information in a more compensatory way

(Bröder & Eichler, 2006; Newell & Fernandez, 2006; Newell & Shanks, 2004; Pohl, 2006; Richter & Späth, 2006)

It seems that the use of a relatively simple heuristic such as the recognition heuristic is integrated with other relevant knowledge (Richter & Späth, 2006), so much so that people choose to buy more information than they strictly need, even in relatively simple decision making tasks (Newell & Shanks, 2003). However, the gathering of additional information is sensitive to the costs associated with obtaining more information (Bröder, 2000), and with practice those cues that are selected are those that have previously been useful in making a successful decision (Newell, Rakow, Weston & Shanks, 2004). That is, the costs and benefits of obtaining more information seems to be taken into account, suggesting cue selection is adaptive to the task constraints (Bröder, 2000).

Thus there is a body of prior research suggesting cue selection is often compensatory and makes use of more than a single cue in coming to a decision. Does this extend to more real-world lie-truth judgments? My research suggests so. Experiment 4 showed that consistency was a diagnostic indicator of deception and was used by around 150 participants across three experiments. Yet it was found that, despite using a diagnostic cue, raters' accuracy did not reflect this. I suggested raters might have been taking additional information into account besides consistency cues.

If raters take into account a number of different cues, what might these be? A successful decision-maker should select cues that are informative. Yet until recently one of the major explanations of low lie detection accuracy was that participants have the wrong beliefs about which cues are diagnostic of deception (Ekman, 1992; Miller & Stiff, 1993; The Global Deception Team, 2006; Vrij, Granhag & Mann, 2010). Lie detection researchers argued raters were making use of information that was not just non-diagnostic, but would actually suggest the incorrect judgment. More recently, four meta-analyses conducted by Hartwig and Bond (2011) showed that, despite self-reporting use of the wrong cues, raters are not using them. Instead, they rely relatively little on erroneous cues such as eye contact and more on diagnostic indicators. The message from these meta-analyses and my own research is that raters seem to select context-appropriate information. I claim an adaptive decision maker account may prove to be a useful description of how raters are making use of these cues, and that it offers a promising new direction for lie detection research.

There is some suggestive evidence in the lie detection literature indicating a compensatory decision strategy. A recent study by Bond and colleagues (Bond et al., 2013) found raters made use of a cue that was 100% diagnostic, namely the incentive to lie. But when visual cues from the statement became available raters made use of those to their own detriment, suffering a marked drop in accuracy from 97% to 76%. Even when there is a clear 'best' raters do not always take it without regard for other cues (c.f. Gigerenzer & Goldstein, 1996). Additional information can be taken into the lie-truth decision process, even when that information is counterproductive.

This raises the issue of what cues are selected and whether all cues are treated equally. A compensatory strategy, as I am arguing is used in lie detection, requires weighting the available evidence so that its net effect can be determined. Research in decision making suggests salient cues are given heavier weighting in the decision process (Platzer & Bröder, 2012). Taking this to the lie-detection arena, our research paints a similar picture (Street & Richardson, in prep.). In the first experiment, participants rated a set of behaviours for how tense the speaker appeared (TE) and how hard they appeared to be thinking (TH), without any forewarning as to the deceptive nature of the stimuli. In a second experiment, a new set of participants were shown one half of the stimulus set where the two cues that had been rated in the previous experiment, i.e. the TE and TH cues, were in greatest contrast to each other. That is, those statements in which speakers appeared relatively tense but not appearing to think hard, or vice versa, were selected, because these cues suggest competing interpretations of deception and honesty, respectively. Participants in the second study were asked to rate either whether the speaker appeared tense or not, or whether the speaker appeared to be thinking hard or not (within subjects) as the speaker delivered the statement. That is, throughout the course of the statement the rater gave a continuous judgment, as per the methodology employed in Experiment 5. At the end of each video the rater made a lie-truth judgment.

It was found that when raters attended to a cue that suggested deceit (i.e. the speaker appeared to be thinking hard or tense), there was a significant decline in the truth ratings compared to when raters, viewing the exact same behaviour, were attending to a cue that suggested honesty. By making particular cues salient, they came to have a more forceful impact on the final decision outcome. I argue this evidences the greater impact of salient information in making lie-truth judgments, similar to the findings of research in less real-world tasks (Platzer & Bröder, 2012).

To summarise, it seems a dual-process account of the truth bias has difficulties both on the empirical level, as demonstrated in this thesis, as well as on the conceptual level, as illustrated in the preceding chapter. I have suggested a single-process model is both more parsimonious and is able to make (and survive the testing of) novel predictions. Although the data do not allow for any firm conclusions to be drawn about the compensatory or non-compensatory strategies employed, some of the presented evidence is at least consistent with a compensatory strategy, where salient cues have a greater impact on the judgment outcome.

Future Directions

Conflict Between Prior Knowledge and Online Processing

In this thesis I have focused on the cognitive processes underlying the decision process independently of the speakers' behaviours. However, it is of course important to consider how the behaviour is interpreted in making the judgment. For example, if a speaker delivers an implausible story, the rater may be expected to become suspicious and to be more likely to consider the statement a lie. But what is unclear is how raters bring together potentially conflicting information from the speaker, which may suggest deception, and from their prior experience with the world, that people usually tell the truth. Understanding how the two come together and how this potential conflict is resolved is an interesting area for future exploration.

Of the experiments conducted, Experiment 6 may be best set up to see the conflict, where the behaviours were held constant but the beliefs about the base rate were manipulated. Participants rating lies in the high-truth expectancy condition, for example, would have had to deal with deceptive behaviours conflicting with their prior belief about the high base rate of honesty. However, there was no evidence for a conflict between the speaker's behaviour (i.e. whether they lied or told the truth) and base-rate beliefs during the presentation of the behaviour, i.e. as the belief judgment was forming.

Although this could suggest the prior beliefs and the behaviour of the speaker have independent effects on the developing judgment, the lack of conflict between behaviour and prior beliefs may result from the fact that raters were not particularly accurate in detecting deception. The leaked deceptive behaviours may not have had a substantial influence on the developing judgment because they may not have been highly diagnostic or may have been missed by the rater. Consider that both truth-tellers and liars want or expect to be believed by the listener. Self-presentational behaviours may have given an impression of honesty regardless of the veracity of the speaker's statement. Over time, more behaviour will become available from the speaker. This could potentially explain why, over time, raters show an increase in their truth belief regardless of veracity (Experiment 6). That is, while initially the raters' expectations may guide the judgment, as more (self-presentational) behaviour becomes available it may become the basis for making the judgment. This is of course speculative, but is worth exploring in future research.

In Experiment 6, after having viewed the full statement, raters were required to make a single judgment of the statement as either a truth or lie. By this point, after having heard the full statement and seen the full behavioural display, when raters expected mostly lies they showed no response bias whereas those expecting mostly truths showed a strong truth bias. This could have resulted from the interaction between the base rate information provided and the behaviour of the speaker. The lack of bias when expecting mostly lies, for example, may reflect the competing interpretations suggested by the speaker's behaviours, which are self-presentational in an attempt to appear (rightly or wrongly) honest, and the rater's base rate belief that suggested the speaker was likely to be lying.

However, the finding is also consistent with an explanation of competition between the base rate beliefs given to participants in this study (given base rate) and their own base rate beliefs that they have learnt from their interactions outside of the laboratory (learnt base rate). Because people tell the truth more often than they lie (e.g., DePaulo et al., 1996), participants' learnt base rate will suggest the speaker is likely to be telling the truth, regardless of the given base rate of honesty in this study.

To begin exploring the interaction between behaviour and cognition, it may be fruitful to move to a task where the cues in the environment can be selected and controlled, and where prior knowledge of the task is known or minimised. For example, a set of novel creatures could be invented that display three simple behaviours. Raters could be trained to learn whether the presence or absence of a particular combination of behaviours suggests truth telling or lying, with each combination having a defined probability of appearing given truth-telling/lying. The training would serve as a form of prior knowledge on a task where the past experience has been controlled. The behaviours could be easily manipulated in a later test phase.

Lie detection has been the topic of interest here because it offers a somewhat naturalistic environment where uncertainty is high due to little reliability in the cues presented. There are a number of important questions one may ask moving to a similar but more controllable paradigm: under uncertainty, will raters give more weighting to cues that are environment-based or are drawn from their prior experiences? Will this depend on the diagnosticity of these cue-types, how difficult it is to make use of them, or how salient each is? An adaptive system may be expected to satisfice and make use of relatively easily accessible information that has some diagnostic validity to make a 'good enough' decision. It will be important to be able to quantify these variables in order to determine how these factors contribute to the judgment.

Internal and External Uncertainty

A distinction was made between internal and external uncertainty in the discussion of Experiment 7. The claim here was that internal uncertainty reflects indecision in light of the evidence, while external uncertainty reflects a lack of information on which to attempt to make a judgment. The former, internal uncertainty, can be used as a heuristic in a given context: ‘an inability to decide results from a stimulus that is difficult to categorise’. In the current context, lies are more difficult to spot than truths (Levine et al., 1999), and so uncertainty can be used to make an informed ‘lie’ guess. The latter, external uncertainty, cannot be used in this way. The uncertainty is not a result of a difficulty in making the decision, but an absence of information on which to even begin making one. In the absence of evidence, context-relevant knowledge can be used, such as base rate information (Experiment 6). This was the first study to explore the effects of perceived base rates on the lie-truth judgment.

An attempt to test the validity of this distinction was made in Experiment 8, replicating earlier work on the issue (Gilbert et al., 1990, Study 1). The aim here was to show that the data supposedly supporting a dual-process Spinozan mind position was actually reflecting raters’ attempts to make the best guess under uncertainty, whether that uncertainty was internal or external. However, it became clear that the pattern of responding that appeared to support the Spinozan position resulted from a design flaw in the experiment.

Support for the reliance of prior knowledge information under external uncertainty was shown in Experiment 6: given different contextual information (about the base rates), those who had to guess used this base rate information during the early moments of processing. Over time, there was less reliance on the base rate information, such that by the end of the trial when the full statement had been presented, those who expected four out of five speakers to lie nonetheless rated approximately half of them as truth-tellers.

The internal-external uncertainty claim leads to a somewhat unusual prediction. Before the speaker begins delivering their statement, i.e. in the absence of any information, raters should still be inclined to guess the speaker will tell the truth. Context-relevant information should be relied upon to make a guess, and can be made even before the speaker begins delivering their statement. By manipulating the context to one where raters should anticipate people to lie, such as manipulating base rate beliefs or leading participants to believe they are listening to sales people, the guessing strategy should reflect this. Importantly, the Spinozan mind position makes no predictions in this regard, so support for this prediction would not contradict the Spinozan position.

Internal uncertainty is that which is present after having collected some information from the task. If, having evaluated the evidence, a clear decision is not made, this inability to decide can itself be used to guide the judgment. Once again, the key to how this uncertainty is used is context-dependent. In the context of lie detection, lies are harder to detect than truths (Levine et al., 1999), and so should be more likely to result in indecision. But if lies were easier to detect than truths, this effect should reverse: under internal uncertainty, raters should guess 'true' because truths would be more likely to result in indecision.

One way to get at this would be to give participants a small list of behavioural cues written on screen. Each cue could have assigned with it a probability of being present given that the statement is a lie and a second probability of being present given that the statement is a truth. On each trial a subset of these cues could be produced as abstracted descriptions (e.g. 'This person stuttered and told a plausible story') and the rater would need to judge whether the speaker lied or told the truth. If the cues tend to have a high probability of being present if the speaker is lying (i.e.

they are highly diagnostic) but are almost equiprobable to be present if the speaker is telling the truth (i.e. they have relatively low diagnostic validity), it would be expected that raters would guess ‘truth’ when they are unsure. This could be tested by employing the response paradigm used in Experiments 5 and 7, such that half the participants are forced to judge and the other half are not.

The task could be abstracted further, such that the context of deception is removed and participants would need to categorise arbitrary items (e.g. polygons) into one of two nonsense categories (e.g. squibbers or cringers) based on a set of cues, each of which has an associated probability of being present given that the item belongs to the squibbers or cringers category. This would reduce the influence of prior knowledge such as base rate information that may influence a lie-detection decision.

Ultimately, how decisions are made under uncertainty is thought to be dependent on contextual information and prior knowledge. Manipulation of these should influence the form of guessing that respondents make.

Finally, the social orientation experiments in Chapter 6 began to explore the role of social information in the decision-making task. I suggested the data show no evidence of a willingness to sacrifice accuracy for the sake of abiding by social rules, but rather that social information can be used as an additional cue to guide the decision. Whether this information comes in the form of a perceived social attachment with another, evoking feelings of intimacy, or whether it is used as a gauge of the probability with which ingroup members will lie, is unclear. One way to make this distinction would be to manipulate the base rate information, as per Experiment 6, and determine whether the ingroup bias persists.

In summary, future work may seek to discover how behaviour and cognition interplay in forming decisions under uncertainty. Perhaps the most interesting future direction is to explore the context-dependency of decision-making under uncertainty.

Conclusions

This thesis set out to determine how and why raters make systematic decisions under uncertainty. To do so, I employed a real-life socially oriented decision made by most people, from suspicious spouses to cynical customs officials: lie detection. Despite strong prior beliefs and extensive experience (see DePaulo & Kashy, 1998; DePaulo et al., 1996; Hartwig & Bond, 2011; Park et al., 2002; The Global Deception Team, 2006; Vrij, Granhag & Porter, 2010), we are poor lie detectors (Bond & DePaulo, 2006; Kraut, 1980). In this task, we bring information into the task from prior experience with making similar judgments, with similar people, and in similar situations. This ranges from probabilities with which people in general usually lie to a more fine-grained understanding of the social situation and inferences made about the individual speaker, such as whether they are an ingroup member. The rater does not rely on this information alone: they take into account the behaviours of the speaker as they unfold (see Burgoon & Buller, 1996). Yet in spite of this rich source of information raters are largely inaccurate and uncertain, making lie detection an interesting test case for decision-making under uncertainty.

Dual process models are currently gaining popularity in the lie detection field. On the surface there are a number of phenomena that appear to show the effects of dual processes (Gilbert, 1991; Gilbert et al., 1990; Gilbert et al., 1993; Masip et al., 2006; Masip, Garrido, et al., 2009; Masip et al., 2010; Reinhard & Sporer, 2010; Reinhard, Sporer & Marksteiner, 2008). In a series of experiments the predictions made by these theories were tested. I found that the heuristic-analytic model failed to meet the challenges issued against it.

Instead I have shown that the data are consistent with a context-dependent adaptive decision-maker, one that makes the best guess in an information-limited world. In support of this view, once the context was changed, raters adapted to it. When consistency information was available, for example, it was employed. The same was true of social relatedness information and base rate knowledge. This information was used to make an informed guess under uncertainty, as shown in Chapter 5.

This thesis concludes that we are adaptive decision makers, balancing the demands of the task and the availability of information with the limited cognitive resources available (Gigerenzer et al., 1999; Payne et al., 1993; Simon, 1990; Stewart, Chater & Brown, 2006). To quote Fiske (1992, p. 879), 'people are no fools; they use workable strategies with adequate outcomes'.

References

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *The Forensic Examiner, 15*, 6-11.

Ajzen, I., & Sexton, J. (1999). Depth of processing, belief congruence, and attitude-behavior correspondence. In S. Chaiken & Y. Trope (Eds.), *Dual-Process Theories in Social Psychology* (pp. 117-138). New York: Guilford Press.

Albrechtsen, J. S., Meissner, C. A., & Susa, K. J. (2009). Can intuition improve deception detection performance? *Journal of Experimental Social Psychology, 45*(4), 1052-1055.

Allen, R. B., & Ebbesen, E. B. (1981). Cognitive processes in person perception: Retrieval of personality trait and behavioral information. *Journal of Experimental Social Psychology, 17*(2), 119-141.

Allen, V. L., & Wilder, D. A. (1979). Group categorization and attribution of belief similarity. *Small Group Behavior, 10*(1), 73-80.

Allport, G. W. (1954). *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley.

Allyn, J., & Festinger, L. (1961). The effectiveness of unanticipated persuasive communications. *Journal of Abnormal and Social Psychology, 62*(1), 35-40.

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review, 13*(3), 219-239.

Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language, 31*, 685-712.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*(2), 256-274.

Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology, 64*(3), 431-441.

Anderson, D. E., DePaulo, B. M., & Ansfield, M. E. (2002). The development of deception detection skill: A longitudinal study of same-sex friends. *Personality and Social Psychology Bulletin, 28*(4), 536-545.

Anderson, S. E., Huette, S., Matlock, T., & Spivey, M. (2009). On the temporal dynamics of negated perceptual simulations. In F. Paril & V. Tobin & M. Turner (Eds.), *Meaning, Form & Body*. Stanford, CA: CSLI.

Anderson, S. E., Huette, S., Matlock, T., & Spivey, M. (2010). Comprehending negated sentences with binary states and locations. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Argyle, M., & Henderson, M. (1984). The rules of friendship. *Journal of Social and Personal Relationships, 1*, 211-237.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 1230-1240.

Aschenbrenner, K. M., & Schaefer, R. E. (1980). Minimal group situations: Comments on a mathematical model and on the research paradigm. *European Journal of Social Psychology, 10*, 389-398.

Ask, K., & Granhag, P. A. (2007). Motivational bias in criminal investigators' judgments of witness reliability. *Journal of Applied Social Psychology, 37*(3), 561-591.

Ask, K., Rebelius, A., & Granhag, P. A. (2008). The 'elasticity' of criminal evidence: A moderator of investigator bias. *Applied Cognitive Psychology, 22*, 1245-1259.

Aune, R. K., Metts, S., & Hubbard, A. S. E. (1998). Managing the outcomes of discovered deception. *The Journal of Social Psychology, 138*(6), 677-689.

Ayton, P., Hunt, A. J., & Wright, G. (1989). Psychological conceptions of randomness. *Journal of Behavioral Decision Making, 2*, 221-238.

Baayen, R. H. (2008). *Analysing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge, UK: Cambridge University Press.

Backbier, E., Hoogstraten, J., & Terwogt-Kouwenhoven, K. M. (1997). Situational determinants of the acceptability of telling lies. *Journal of Applied Social Psychology, 27*(12), 1048-1062.

Ball, L. J., 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.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica, 44*, 211-233.

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From statistical formats to cognitive structures. *Behavioral and Brain Sciences, 30*(3), 287-297.

Bargh, J. A. (1989). Conditional automaticity: Variety of automatic influences in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended Thought* (pp. 3-51). New York: Guilford Press.

Baron, R. S., Baron, P. H., & Miller, N. (1973). The relation between distraction and persuasion. *Psychological Bulletin, 80*(4), 310-323.

Bates, D., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes. R package version.

Bauchner, J. E., Brandt, D. R., & Miller, G. R. (1977). The truth/deception attribution: Effects of varying levels of information availability. In B. R. Ruben (Ed.), *Communication Yearbook 1* (pp. 229-243). New Brunswik, NJ: Transaction Books.

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review, 3*, 439-449.

Beach, L. R., & Swensson, R. G. (1967). Instructions about randomness and run dependency in two-choice learning. *Journal of Experimental Psychology, 75*, 279-282.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General, 121*(4), 446-458.

Bell, K. L., & DePaulo, B. M. (1996). Liking and lying. *Basic and Applied Social Psychology, 18*(3), 243-266.

Benoit, W. L. (1998). Forewarning and persuasion. In M. Allen & R. Priess (Eds.), *Persuasion: Advances Through Meta-analysis* (pp. 159-184). Cresskill, NJ: Hampton Press.

Beukeboom, C. J., Finkenauer, C., & Wigboldus, D. H. J. (2010). The negation bias: When negations signal stereotypic expectancies. *Journal of Personality and Social Psychology, 99*(6), 978-992.

Billig, M., & Tajfel, H. (1973). Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology, 3*(1), 27-52.

Blair, J. P. (2006). From the field: Can detection of deception response bias be manipulated? *Journal of Crime and Justice, 29*(2), 141-152.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review, 113*(4), 700-765.

Bohner, G., Moskowitz, G. B., & Chaiken, S. (1995). The interplay of heuristic and systematic processing of social information. *European Review of Social Psychology, 6*(1), 33-68.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214-234.

Bond, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*(4), 477-492.

Bond, C. F., Howard, A. R., Hutchison, J. L., & Masip, J. (2013). Overlooking the obvious: Incentives to lie. *Basic and Applied Social Psychology, 35*, 212-221.

Bond, C. F., Omar, A., Pitre, U., Lashley, B. R., Skaggs, L. M., & Kirk, C. T. (1992). Fishy-looking liars: Deception judgment from expectancy violation. *Journal of Personality and Social Psychology, 63*(6), 969-977.

Bond, G. D., Malloy, D. M., Arias, E. A., Nunn, S. N., & Thompson, L. A. (2005). Lie-based decision making in prison. *Communication Reports, 18*(1-2), 9-19.

Boon, S. D., & McLeod, B. A. (2001). Deception in romantic relationships: Subjective estimates of success at deceiving and attitudes toward deception. *Journal of Social and Personal Relationships, 18*(4), 463-476.

Borkenau, P., Mauer, N., Rieman, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Personality Processes and Individual Differences, 86*(4), 599-614.

Boyes, A. D., & Fletcher, G. J. O. (2007). Meta-perceptions of bias in intimate relationships. *Journal of Personality and Social Psychology, 92*, 286-306.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436.

Bransford, J. D., & Johnson, M. K. (1973). Considerations of some problems of comprehension. In W. Chase (Ed.), *Visual Information Processing* (pp. 383-438). New York: Academic Press.

Brewer, M. B. (1979). Ingroup bias in the minimal intergroup situation: Cognitive-motivational analysis. *Psychological Bulletin, 86*(2), 307-324.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55*(3), 429-444.

Brewer, M. B., & Silver, M. (1978). Ingroup bias as a function of task characteristics. *European Journal of Social Psychology, 8*, 393-400.

Brocas, I., & Carrillo, J. D. (in press). Dual-process theories of decision-making: A selective survey. *Journal of Economic Psychology*.

Bröder, A. (2000). Assessing the empirical validity of the "Take-the-Best" heuristic as a model of human proabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1332-1346.

Bröder, A. (2003). Decision making with the "Adaptive Toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 611-625.

Bröder, A., & Eichler, A. (2006). The use of recognition information and additional cues in inferences from memory. *Acta Psychologica, 121*, 275-284.

Bröder, A., & Schiffer, S. (2003). Take the best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General, 132*(2), 277-293.

Brown, R. (2000). Social identity theory: Past achievements, current problems and future challenges. *European Journal of Social Psychology, 30*(6), 745-778.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203-242.

Burge, T. (1993). Content preservation. *Philosophical Review, 101*, 457-488.

Burgoon, J. K., Blair, J. P., & Strom, R. E. (2008). Cognitive biases and nonverbal cue availability in detecting deception. *Human Communication Research, 34*, 572-599.

Burgoon, J. K., Stoner, G. M., Bonito, J. A., & Dunbar, N. E. (2003). Trust and deception in mediated communication, *Proceedings of the 36th Hawaii International Conference on System Sciences*. Hawaii.

Burgoon, J. K., & Walther, J. B. (1990). Nonverbal expectancies and the evaluative consequences of violations. *Human Communication Research, 17*(2), 232-265.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic cognitive approach to decision-making in an uncertain environment. *Psychological Review, 100*(3), 432-459.

Carlston, D. E., & Skowronski, J. J. (1986). Trait memory and behavior memory - the effects of alternative pathways on impression judgment response-times. *Journal of Personality and Social Psychology, 50*(1), 5-13.

Carney, D. R., Colvin, C. R., & Hall, J. A. (2007). A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality, 41*, 1054-1072.

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review, 99*, 122-149.

Caspi, A., & Gorsky, P. (2006). Online deception: Prevalence, motivation, and emotion. *CyberPsychology & Behavior, 9*(1), 54-59.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39*(5), 752-766.

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna & J. M. Olson & C. P. Herman (Eds.), *Social Influence: The Ontario Symposium* (Vol. 5, pp. 3-39). Mahwah, NJ: Lawrence Erlbaum Associates.

Chaiken, S., & Eagly, A. H. (1983). Communication modality as a determinant of persuasion: The role of communicator salience. *Journal of Personality and Social Psychology, 45*, 221-256.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended Thought* (pp. 212-252). New York: Guilford Press.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology, 66*(3), 460-473.

Chaiken, S., & Stangor, C. (1987). Attitudes and attitude change. *Annual Review of Psychology, 38*, 575-630.

Chaiken, S., & Trope, Y. (1999). *Dual-Process Theories in Social Psychology*. New York: Guilford Press.

Chase, V. M., Hertwig, R., & Gigerenzer, G. (1998). Visions of rationality. *Trends in Cognitive Sciences, 2*(6), 206-214.

Chen, S., & Chaiken, S. (1999). The Heuristic-Systematic Model in its Broader Context. In S. Chaiken & Y. Trope (Eds.), *Dual-Process Theories in Social Psychology* (pp. 73-96). New York: Guilford Press.

Chen, S., Shechter, D., & Chaiken, S. (1996). Getting at the truth or getting along: Accuracy- versus impression-motivated heuristic and systematic processing. *Journal of Personality and Social Psychology, 71*(2), 262-275.

Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review, 99*(1), 431-457.

Chen, Y. W., & Blanchard-Fields, F. (2000). Unwanted thought: Age differences in the correction of social judgments. *Psychology and Aging, 15*, 475-482.

Chun, W. Y., & Kruglanski, A. W. (2006). The role of task demands and processing resources in the use of base rate and individuating information. *Journal of Personality and Social Psychology, 91*, 205-217.

Chung, S., & Fink, E. L. (2008). The cognitive dynamics of beliefs: The effect of information on message processing. *Human Communication Research, 34*(3), 477-504.

Cialdini, R. B., & Petty, R. E. (1981). Anticipatory opinion effects. In R. E. Petty & T. M. Ostrom & T. C. Brock (Eds.), *Cognitive Responses in Persuasion* (pp. 217-235). Hillsdale, NJ: Erlbaum.

Clark, H. H., & Chase, W. G. (1972). Process of comparing sentences against pictures. *Cognitive Psychology, 3*(3), 472-517.

Clark, H. H., & Clark, E. V. (1977). *Psychology and Language: An Introduction to Psycholinguistics*. New York: Harcourt Brace Jovanovich.

Clark, M. S., & Lemay, E. P. (2010). Close relationships. In S. T. Fiske & D. T. Gilbert & G. Lindzey (Eds.), *Handbook of Social Psychology* (5 ed., Vol. 2, pp. 898-940). Hoboken, NJ: John Wiley & Sons, Inc.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed. ed.). Hillsdale, N.J.: L. Erlbaum Associates.

Cole, T. (2001). Lying to the one you love: The use of deception in romantic relationships. *Journal of Social and Personal Relationships, 18*, 107-129.

Colwell, L. H., Colwell, K., Hiscock-Anisman, C. K., Hartwig, M., Cole, L., Werdin, K., & Youschak, K. (2012). Teaching professionals to detect deception: The efficacy of a brief training workshop. *Journal of Forensic Psychology Practice, 12*(1), 68-80.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). Accuracy and frugality in a tour of environments. In G. Gigerenzer & P. M. Todd & T. A. Research Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 97-118). New York: Oxford University Press.

De Neys, W. (2006). Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual-task considerations. *Quarterly Journal of Experimental Psychology, 59*, 1070-1100.

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin, 111*, 203-243.

DePaulo, B. M., Ansfield, M. E., Kirkendol, S. E., & Boden, J. M. (2004). Serious lies. *Basic and Applied Social Psychology, 26*(2-3), 147-167.

DePaulo, B. M., & Bell, K. L. (1996). Truth and investment: Lies are told by those who care. *Journal of Personality and Social Psychology, 71*(4), 703-716.

DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review, 1*(4), 346-357.

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology, 74*(1), 63-79.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*(5), 979-995.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74-118.

DePaulo, B. M., & Morris, W. L. (2004). Discerning lies from truths: Behavioural cues to deception and the indirect pathway of intuition. In P. A. Granhag & L. A. Strömwall (Eds.), *The Detection of Deception in Forensic Contexts* (pp. 15-40). Cambridge: Cambridge University Press.

DePaulo, B. M., & Rosenthal, R. (1979). Telling Lies. *Journal of Personality and Social Psychology, 37*(10), 1713-1722.

DePaulo, B. M., Stone, J. I., & Lassiter, G. D. (1985). Deceiving and detecting deceit. In B. R. Schlenker (Ed.), *The Self and Social Life* (pp. 323-370). New York: McGraw-Hill.

DePaulo, B. M., Zuckerman, M., & Rosenthal, R. (1980). Detecting deception: Modality effects. In L. Wheeler (Ed.), *The Review of Personality and Social Psychology* (Vol. 1, pp. 125-162). Beverly Hills, CA: Sage.

DePaulo, P. J., & DePaulo, B. M. (1989). Can deception by salespersons and customers be detected through nonverbal behavioral cues? *Journal of Applied Social Psychology, 19*(18), 1552-1577.

Deutsch, R., Gawronski, B., & Strack, F. (2006). At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology, 91*(3), 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*(6), 434-446.

Development Core Team, R. (2011). R: A language and environment for statistical computing. Vienna, Austria.

Diehl, M. (1989). Justice and discrimination between minimal groups: The limits of equity. *British Journal of Social Psychology, 28*(3), 227-238.

Diehl, M. (1990). The minimal group paradigm: Theoretical explanations and empirical findings. *European Review of Social Psychology, 1*(1), 263-292.

Dodd, D. H., & Bradshaw, J. M. (1980). Leading questions and memory: Pragmatic constraints. *Journal of Verbal Learning & Verbal Behavior, 19*, 695-704.

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, *121*(3), 275-277.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30*, 412-431.

Duran, N. D., Dale, R., Kello, C. T., Street, C. N. H., & Richardson, D. C. (2013). Exploring the movement dynamics of deception. *Frontiers in Cognitive Science, 4*(10), 1-16.

Echterhoff, G., Higgins, E. T., & Levine, J. M. (2009). Shared reality: Experiencing commonality with others' inner states about the world. *Perspectives on Psychological Science, 4*(5), 496-521.

Ekman, P. (1992). *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*. New York: W. W. Norton & Co.

Ekman, P., Friesen, W. V., O'Sullivan, M., & Scherer, K. R. (1980). Relative importance of face, body, and speech in judgments of personality and affect. *Journal of Personality and Social Psychology, 38*(2), 270-277.

Elaad, E. (2000). Detection of deception. In J. A. Siegel & P. J. Saukko & G. C. Knupfer (Eds.), *Encyclopedia of Forensic Sciences* (pp. 550-556). London: Academic Press.

Elaad, E. (2003). Effects of feedback on the overestimated capacity to detect lies and the underestimated ability to tell lies. *Applied Cognitive Psychology, 17*, 349-363.

Erb, H. P., Kruglanski, A. W., Chun, W. Y., Pierro, A., Mannetti, L., & Spiegel, S. (2003). Searching or commonalities in human judgment: The parametric unimodel and its dual mode alternatives. *European Review of Social Psychology, 14*, 1-47.

Estes, W. K. (1964). Probability learning. In A. W. Melton (Ed.), *Categories of Human Learning* (pp. 89-128). New York: Academic Press.

Evans, J. S. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review, 13*, 378-395.

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

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

Evans, J. S. 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.

Evans, J. S. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human Reasoning: The Psychology of Deduction*. Hove: Lawrence Erlbaum Associates Ltd.

Evans, J. S. B. T., & Stanovich, K. E. (2013). Theory and metatheory in the study of dual processing: Reply to comments. *Perspectives on Psychological Science, 8*(3), 263-271.

Fan, R. M. T., Wagner, H. L., & Manstead, A. S. R. (1995). Anchoring, familiarity, and confidence in the detection of deception. *Basic and Applied Social Psychology, 17*(1-2), 83-96.

Fenn, E., Newman, E. J., Pezdek, K., & Garry, M. (2013). The effect of nonprobative photographs on truthiness persists over time. *Acta Psychologica, 144*(1), 207-211.

Fiedler, K., Armbruster, T., Nickel, S., Walther, E., & Asbeck, J. (1996). Constructive biases in social judgment: Experiments on the self-verification of question contents. *Journal of Personality and Social Psychology, 71*(5), 861-873.

Fiedler, K., & Walka, I. (1993). Training lie-detectors to use nonverbal cues instead of global heuristics. *Human Communication Research, 20*(2), 199-223.

Fiori, M. (2009). A new look at emotional intelligence: A dual-process framework. *Personality and Social Psychology Review, 13*(1), 21-44.

Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from daguerrotype to laserphoto. *Journal of Personality and Social Psychology, 63*(6), 877-889.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum model of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 23, pp. 1-74). New York: Academic Process.

Fiske, S. T., & Taylor, S. E. (1991). *Social Cognition* (2nd ed. ed.). New York: McGraw-Hill.

Fitzsimons, G. J., & Shiv, B. (2001). Nonconscious and contaminative effects of hypothetical questions on subsequent decision making. *Journal of Consumer Research, 28*(2), 224-238.

Fleming, J., & Arrowood, A. J. (1979). Information processing and the perseverance of discredited self-perceptions. *Personality and Social Psychology Bulletin, 5*, 201-205.

Forrest, J. A., & Feldman, R. S. (2000). Detecting deception and judge's involvement: Lower task involvement leads to better lie detection. *Personality and Social Psychology Bulletin, 26*(1), 118-125.

Fowler, K. A., Lilienfeld, S. O., & Patrick, C. J. (2009). Detecting psychopathy from thin slices of behavior. *Psychological Assessment, 21*(1), 68-78.

Fox, R. C. (1957). Training for uncertainty. In R. K. Merton & C. G. Reader & P. L. Kendall (Eds.), *The Student-Physician: Introductory Studies in the Sociology of Medical Education* (pp. 207-241). Cambridge, MA: Harvard University Press.

Frank, M. G., & Feeley, T. H. (2003). To catch a liar: Challenges for research in lie detection training. *Journal of Applied Communication Research, 31*(1), 58-75.

Fraundorf, S. H., Benjamin, A. S., & Watson, D. G. (2013). What happened (and what did not): Discourse constraints on encoding of plausible alternatives. *Journal of Memory and Language, 69*(3), 196-227.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178-210.

Funder, D. C. (1996). Base rates, stereotypes, and judgmental accuracy. *Behavioral and Brain Sciences, 19*, 22-23.

Gaertner, S. L., & Dovidio, J. F. (2000). *Reducing intergroup bias: The common ingroup identity model*. Philadelphia: Psychology Press.

Gaissmaier, W., Straubinger, N., & Funder, D. C. (2007). Ecologically structured information: The power of pictures and other effective data presentations. *Behavioral and Brain Sciences, 30*(3), 263-264.

Gawronski, B., Sherman, J. W., & Trope, Y. (in press). Two of what? A conceptual analysis of dual-process theories. In J. W. Sherman & B. Gawronski & Y. Trope (Eds.), *Dual-Process Theories of the Social Mind*. New York: Guilford Press.

George, J. F., & Robb, A. (2008). Deception and computer-mediated communication in daily life. *Communication Reports, 21*, 92-103.

Gigerenzer, G. (2007). *Gut Feelings: Shortcuts to Better Decision Making*. London: Penguin.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*(1), 107-143.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451-482.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4), 650-669.

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: Take the best and its relative. In G. Gigerenzer & P. M. Todd & T. A. Research Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 75-96). New York: Oxford University Press.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98*, 506-528.

Gigerenzer, G., Martignon, L., Hoffrage, U., Rieskamp, J., & Czerlinski, J. (2008). One-reason decision making. In C. R. Plott & V. L. Smith (Eds.), *Handbook of Experimental Economics Results* (Vol. 1, pp. 1004-1017). Amsterdam: Elsevier.

Gigerenzer, G., & Regier, T. (1996). How do we tell an association from a rule? Comment on Sloman (1996). *Psychological Bulletin, 119*, 23-26.

Gigerenzer, G., & Selten, R. (Eds.). (2001). *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.

Gigerenzer, G., Todd, P. M., & The ABC Research Group (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46*(2), 107-119.

Gilbert, D. T., & Krull, D. S. (1988). Seeing less and knowing more: The benefits of perceptual ignorance. *Journal of Personality and Social Psychology, 54*(2), 193-202.

Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology, 59*(4), 601-613.

Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology, 65*(2), 221-233.

Gilovich, T., Griffin, D. W., & Kahneman, D. (2002). *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge: Cambridge University Press.

Gilovich, T., Savitsky, K., & Medvec, V. H. (1998). The illusion of transparency: biased assessments of others' ability to read one's emotional states. *Journal of personality and social psychology, 75*(2), 332-346.

Ginossar, Z., & Trope, Y. (1980). The effects of base rates and individuating information on judgments about another person. *Journal of Experimental Social Psychology, 16*, 228-242.

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes, 28*(1), 1-26.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language, 43*(3), 379-401.

Glenberg, A. M., Robertson, D. A., Jansen, J. L., & Johnson-Glenberg, M. C. (1999). Not propositions. *Journal of Cognitive Systems Research, 1*, 19-33.

Goldstein, D. G., & Gigerenzer, G. (2008). The recognition heuristic and the less-is-more effect. In C. R. Plott & V. L. Smith (Eds.), *Handbook of Experimental Economics Results* (Vol. 1, pp. 987-992). Amsterdam: Elsevier.

Granhag, P. A. (2006). Rethinking implicit lie detection. *The Journal of Credibility Assessment and Witness Psychology, 7*(3), 180-190.

Granhag, P. A., & Strömwall, L. A. (1999). Repeated interrogations: Stretching the deception detection paradigm. *Expert Evidence, 7*(3), 163-174.

Granhag, P. A., & Strömwall, L. A. (2000a). Effects of preconceptions on deception detection and new answers as to why lie-catchers often fail. *Psychology, Crime & Law, 6*, 197-218.

Granhag, P. A., & Strömwall, L. A. (2000b). "Let's go over this again…": Effects of repeated interrogations on deception detection performance. In A. Czerederecka & T. Jaskiewicz-Obydzinska & J. Wojcikiewicz (Eds.), *Forensic Psychology and Law: Traditional Questions and New Ideas* (pp. 191-196). Krakow, Poland: Institute of Forensic Research Publishers.

Granhag, P. A., & Strömwall, L. A. (2001a). Deception detection based on repeated interrogations. *Legal and Criminological Psychology, 6*(1), 85-101.

Granhag, P. A., & Strömwall, L. A. (2001b). Deception detection: Interrogators' and observers' decoding of consecutive statements. *Journal of Psychology, 135*(6), 603-620.

Granhag, P. A., & Strömwall, L. A. (2002). Repeated interrogations: Verbal and non-verbal cues to deception. *Applied Cognitive Psychology, 16*, 243-257.

Grant, P. R., & Brown, R. (1995). From ethnocentrism to collective protest: Responses to relative deprivation and threats to social identity. *Social Psychology Quarterly, 58*(3), 195-212.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*(5), 701-721.

Greuel, L. (1992). Police officers' beliefs about cues associated with deception in rape cases. In F. Lösel & D. Bender & T. Bliesener (Eds.), *Psychology and Law - International Perspectives* (pp. 234-239). Berlin: Walter de Gruyter.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics* (Vol. 3, pp. 41-58). New York: Academic Press.

Griffin, R. J., Neuwirth, K., Geise, J., & Dunwoody, S. (2002). Linking the heuristic-systematic model and depth of processing. *Communication Research, 29*(6), 705-732.

Guthrie, J., & Kunkel, A. (2013). Tell me sweet (and not-so-sweet) little lies: Deception in romantic relationships. *Communication Studies, 64*(2), 141-157.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814-834.

Hall, K. H. (2002). Reviewing intuitive decision-making and uncertainty: the implications for medical education. *Medical Education, 36*(3), 216-224.

Hancock, J. T., Thom-Santelli, J., & Ritchie, T. (2004). Deception and design: The impact of communication technology on lying behavior. *CHI Letters, 6*(1), 129-134.

Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., & Shadlen, M. N. (2011). Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *Journal of Neuroscience, 31*(17), 6339-6352.

Hardin, G. (1968). The tragedy of the commons. *Science, 162*(3859), 1243-1248.

Harley, T. A. (2008). *The Psychology of Language: From Data to Theory* (Vol. 3). New York: Sage.

Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*(4), 643-659.

Hartwig, M., Granhag, P. A., & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology Crime & Law, 13*(2), 213-227.

Hasson, U., Simmons, J. P., & Todorov, A. (2005). Believe it or not: On the possibility of suspending belief. *Psychological Science, 16*(7), 566-571.

Hawkins, S. A., Hoch, S. J., & Meyers Levy, J. (2001). Low involvement learning: Repetition and coherence in familiarity and belief. *Journal of Consumer Psychology, 11*, 1-11.

Heuttig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*(2), 151-171.

Higgins, E. T., & Bargh, J. A. (1987). Social cognition and social perception. *Annual Review of Psychology, 38*, 369-425.

Hilbig, B. E., & Richter, T. (2011). Homo heuristicus outnumbered: Comment on Gigerenzer and Brighton (2009). *Topics in Cognitive Science, 3*(1), 187-196.

Hill, C., Memon, A., & McGeorge, P. (2008). The role of confirmation bias in suspect interviews: A systematic evaluation. *Legal and Criminological Psychology, 13*(2), 357-371.

Hilton, D. J., & Slugoski, B. R. (2001). Conversational processes in reasoning and explanation. In A. Tesser & N. Schwarz (Eds.), *Blackwell Handbook of Social Psychology: Intrapersonal Processes* (Vol. 1, pp. 181-206). Oxford: Blackwell.

Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology, 47*, 237-271.

Hocking, J. E., Bauchner, J. E., Kaminski, E. P., & Miller, G. R. (1979). Detecting deceptive communication from verbal, visual, and paralinguistic cues. *Human Communication Research, 6*, 33-46.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology, 24*(1), 1-55.

Houston, D. A., & Fazio, R. H. (1989). Biased processing as a function of attitude accessibility: Making objective judgments subjectively. *Social Cognition, 7*(1), 51-66.

Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. D. (1949). *Experiments on Mass Communication*. Princeton: Princeton University Press.

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly, 15*(4), 635-540.

Howard-Pitney, B., Borgida, E., & Omoto, A. M. (1986). Personal involvement: An examination of processing differences. *Social Cognition, 4*, 39-57.

Hurd, K., & Noller, P. (1988). Decoding deception: A look at the process. *Journal of Nonverbal Behavior, 12*(3), 217-233.

Hutchinson, J. M. C., & Gigerenzer, G. (2005). Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. *Behavioural Processes, 69*, 87-124.

Jarvick, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology, 41*, 291-297.

Jeffreys, H. (1961). *Theory of Probability* (3rd ed.). New York: Oxford University Press.

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inferences, and Consciousness*. Cambridge, UK: Cambridge University Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kahlor, L., Dunwoody, S., Griffin, R. J., Neuwirth, K., & Giese, J. (2003). Studying heuristic-systematic processing of risk communication. *Risk Analysis, 23*(2), 355-368.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review, 93*(5), 1449-1475.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*(4), 237-251.

Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist, 60*(3), 215-228.

Kassin, S. M., & Fong, C. T. (1999). "I'm innocent!": Effects of training on judgments of truth and deception in the interrogation room. *Law and Human Behavior, 23*(5), 499-516.

Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior, 27*(2), 187-203.

Kassin, S. M., Meissner, C. A., & Norwick, R. J. (2005). "I'd know a false confession if I saw one": A comparative study of college students and police investigators. *Law and Human Behavior, 29*(2), 211-227.

Keren, G. (2013). A tale of two systems: A scientific advance or a theoretical stone soup? Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science, 8*(3), 257-262.

Keren, G., & Schul, Y. (2009). Two is not always better than one: A critical evaluation of two-system theories. *Perspectives on Psychological Science, 4*(6), 533-550.

Khader, P. H., Pachur, T., Meier, S., Bien, S., Jost, K., & Rösler, F. (2011). Memory-based decision-making with heuristics involves increased activation of decision-relevant memory representations. *Journal of Cognitive Neuroscience, 23*, 3540-3554.

Kiesler, C. A., & Kiesler, S. B. (1964). Role of forewarning in persuasive communications. *Journal of Abnormal and Social Psychology, 68*(5), 547-549.

Kim, R. K., & Levine, T. R. (2011). The effect of suspicion on deception detection accuracy: Optimal level or opposing effects? *Communication Reports, 24*(2), 51-62.

Knowles, E. S., & Condon, C. A. (1999). Why people say "yes": A dual-process theory of acquiescence. *Journal of Personality and Social Psychology, 77*, 379-386.

Koch, A. S., & Forgas, J. P. (2012). Feeling good and feeling truth: The interactive effects of mood and processing fluency on truth judgments. *Journal of Experimental Social Psychology, 48*, 481-485.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, nromative and methodological challenges. *Behavioral and Brain Sciences, 19*, 1-53.

Kraut, R. (1980). Humans as lie detectors: Some second thoughts. *Journal of Communication, 30*(4), 209-216.

Kruglanski, A. W. (1989). The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin, 106*, 395-409.

Kruglanski, A. W., Chen, X., Pierro, A., Mannetti, L., Erb, H. P., & Spiegel, S. (2006). Persuasion according to the unimodel: Implication for cancer communication. *Journal of Communication, 56*, 105-122.

Kruglanski, A. W., Pierro, A., Mannetti, L., Erb, H. P., & Chun, W. Y. (2007). On the parameters of human judgment. *Advances in Experimental Social Psychology, 39*, 225-303.

Kruglanski, A. W., & Thompson, E. P. (1999a). The illusory second mode or, the cue is the message. *Psychological Inquiry, 10*, 182-193.

Kruglanski, A. W., & Thompson, E. P. (1999b). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry, 10*, 83-109.

Kyonka, E. G. E., & Church, R. M. (2005). Empirical tests of "Take-the-Best" with non-human subjects. *Behavioural Processes, 69*, 133-135.

Lane, S. M., & Vieira, K. M. (2012). Steering a new course for deception detection research. *Journal of Applied Research in Memory and Cognition, 1*(2), 136-138.

Lemay, E. P., & Clark, A. G. (2008). You're just saying that: Contingencies of self-worth, suspicion, and authenticity in the partner affirmation process. *Journal of Experimental Social Psychology, 44*, 1376-1382.

Leonardelli, G. J., & Brewer, M. B. (2001). Minority and majority discrimination: When and why. *Journal of Experimental Social Psychology, 37*(6), 468-485.

Levine, T. R. (2010). A few transparent liars. In C. T. Salmon (Ed.), *Communcation Yearbook 34* (pp. 41-61). New York: Sage.

Levine, T. R., Anders, L. N., Banas, J., Baum, K. L., Endo, K., Hu, A. D. S., & Wong, N. C. H. (2000). Norms, expectations, and deception: A norm violation model of veracity judgments. *Communication Monographs, 67*(2), 123-137.

Levine, T. R., Kim, R. K., & Hamel, L. M. (2010). People lie for a reason: Three experiments documenting the principle of veracity. *Communication Research Reports, 27*(4), 271-285.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communication Monographs, 66*(2), 125-144.

Lewis, M. (1993). The development of deception. In M. Lewis & C. Saarni (Eds.), *Lying and Deception in Everyday Life* (pp. 289-296). New York: Guilford Press.

Lopes, L. L., & Oden, G. (1987). Distinguishing between random and nonrandom events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 392-400.

Maier, N. R. F., & Thurber, J. A. (1968). Accuracy of judgments of deception when an interview is watched, heard, and read. *Personnel Psychology, 21*, 23-30.

Marewski, J. N., Gaissmaier, W., & Gigerenzer, G. (2010). Good judgments do not require complex cognition. *Cognitive Processes, 11*, 103-121.

Marr, D. (1982). *Vision: A Computational Investigation in the Human Representation and Processing of Visual Information*. New York: Freeman.

Martin, R., & Hewstone, M. (2003). Majority versus minority influence: When, not whether, source status instigates heuristic or systematic processing. *European Journal of Social Psychology, 33*, 313-330.

Masip, J., Alonso, H., Garrido, E., & Antón, C. (2005). Generalized communicative suspicion (GCS) among police officers: Accounting for the investigator bias effect. *Journal of Applied Social Psychology, 35*(5), 1046-1066.

Masip, J., Alonso, H., Garrido, E., & Herrero, C. (2009). Training to detect what? The biasing effects of training on veracity judgments. *Applied Cognitive Psychology, 23*(9), 1282-1296.

Masip, J., Garrido, E., & Herrero, C. (2006). Observers' decision moment in deception detection experiments: Its impact on judgment, accuracy, and confidence. *International Journal of Psychology, 41*(4), 304-319.

Masip, J., Garrido, E., & Herrero, C. (2009). Heuristic versus systematic processing of information in detecting deception: questioning the truth bias. *Psychological Reports, 105*(1), 11-36.

Masip, J., Garrido, E., & Herrero, C. (2010). Regression toward the mean or heuristic processing in detecting deception?: Reply to Elaad (2010). *Psychological Reports, 107*(2), 587-592.

Mayo, R., Schul, Y., & Burnstein, E. (2004). "I am not guilty" vs "I am innocent": Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology, 40*(4), 433-449.

McCornack, S. A., & Levine, T. R. (1990). When lovers become leery: The relationship between suspicion and accuracy in detecting deception. *Communication Monographs, 57*(3), 219-230.

McCornack, S. A., & Parks, M. R. (1986). Deception detection and relationship development: The other side of trust. In M. L. McLaughlin (Ed.), *Communication Yearbook 9* (pp. 377-389). Beverly Hills, CA: Sage.

McCornack, S. A., & Parks, M. R. (1990). What women know that men don't: Sex differences in determining the truth behind deceptive messages. *Journal of Social and Personal Relationships, 7*(1), 107-118.

McGuire, W. J. (1964). Inducing resistance to persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 1, pp. 191-229). New York: Academic Press.

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior, 26*(5), 469-480.

Merton, R. K. (1948). The self-fulfilling prophecy. *The Antioch Review, 8*(2), 193-210.

Metts, S. (1989). An exploratory investigation of deception in close relationships. *Journal of Social and Personal Relationships, 6*, 159-179.

Millar, M. G., & Millar, K. U. (1997). The effects of cognitive capacity and suspicion on truth bias. *Communication Research, 24*(5), 556-570.

Miller, G. R., Mongeau, P. A., & Sleight, C. (1986). Fudging with friends and lying to lovers: Deceptive communication in personal relationships. *Journal of Social and Personal Relationships, 3*(4), 495-512.

Miller, G. R., & Stiff, J. B. (1993). *Deceptive Communication*. Newbury Park, Calif. ; London: Sage.

Millikan, R. (1987). *Language, Thought and Other Categories*. Cambridge, MA: MIT Press.

Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. L. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the selective reanalysis hypothesis. *Journal of Memory and Language, 59*(3), 266-293.

Moore, S. G., Neal, D. T., Fitzsimons, G. J., & Shiv, B. (2012). Wolves in sheep's clothing: How and when hypothetical questions influence behavior. *Organizational Behavior and Human Decision Processes, 117*(1), 168-178.

Morey, R. D., & Rouder, J. (2013). *BayesFactor: An R package for Bayesian analysis in common research designs*. Available: http://bayesfactorpcl.r-forge.r-project.org [2013, 9 June].

Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods, 16*, 406-419.

Moston, S., Stephenson, G. M., & Williamson, T. M. (1992). The effects of case characteristics on suspect behavior during police questioning. *British Journal of Criminology, 32*(1), 23-40.

Moy, J., & Ng, S. H. (2006). Expectation of outgroup behaviour: Can you trust the outgroup? *European Journal of Social Psychology, 26*, 333-340.

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22*(2), 103-122.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibition-less spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*(3), 226-254.

Newell, B. R., & Fernandez, D. (2006). On the binary quality of recognition and the inconsequentiality of further knowledge: Two critical tests of the recognition heuristic. *Behavioral Decision Making, 19*(4), 333-346.

Newell, B. R., Rakow, T., Weston, N. J., & Shanks, D. R. (2004). Search strategies in decision-making: The success of success. *Journal of Behavioral Decision Making, 17*, 117-137.

Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing "one-reason" decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(1), 53-65.

Newell, B. R., & Shanks, D. R. (2004). On the role of recognition in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 923-935.

Newell, B. R., Weston, N. J., & Shanks, D. R. (2003). Empirical tests of a fast-and-frugal heuristic: Not everyone "takes-the-best". *Organizational Behavior and Human Decision Processes, 91*, 82-96.

Newman, E. J., Garry, M., Bernstein, D. M., Kanter, J., & Lindsay, D. S. (2012). Nonprobative photographs (or words) inflate truthiness. *Psychonomic Bulletin & Review, 19*(5), 969-974.

Newstead, S. E. (2000). Are there two different types of thinking? *Behavioral and Brain Sciences, 23*, 690-691.

Nieuwland, M. S., & Martin, A. E. (2012). If the real world were irrelevant, so to speak: The role of propositional truth-value in counterfactual sentence comprehension. *Cognition, 122*, 102-109.

Nisbett, R. E., & Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgement*. Englewood Cliffs ; London: Prentice-Hall.

O'Sullivan, M. (2003). The fundamental attribution error in detecting deception: The boy-who-cried-wolf effect. *Personality and Social Psychology Bulletin, 29*(10), 1316-1327.

O'Sullivan, M., Ekman, P., & Friesen, W. V. (1988). The effect of comparisons on detecting deceit. *Journal of Nonverbal Behavior, 12*(3), 203-215.

O'Sullivan, M., Ekman, P., Friesen, W. V., & Scherer, K. R. (1985). What you say and how you say it: The contribution of speech content and voice quality to judgments of others. *Journal of Personality and Social Psychology, 48*(1), 54-62.

Oppenheimer, D. M. (2003). Not so fast! (and not so frugal!): Rethinking the recognition heuristic. *Cognition, 90*, B1-B9.

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review, 11*, 988-1010.

Otten, S., & Epstude, K. (2006). Overlapping mental representations of self, ingroup, and outgroup: Unraveling self-stereotyping and self-anchoring. *Personality and Social Psychology Bulletin, 32*(7), 957-969.

Otten, S., & Mummendey, A. (2000). Valence-dependent probability of ingroup-favoritism between minimal groups: An integrative view on the positive-negative asymmetry in social discrimination. In D. Capozza & R. Brown (Eds.), *Social Identity Processes* (pp. 33-48). London: Sage.

Park, H. S., Levine, T. R., McCornack, S. A., Morrison, K., & Ferrara, M. (2002). How people really detect lies. *Communication Monographs, 69*(2), 144-157.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge: Cambridge University Press.

Pennington, N., & Hastie, R. (1991). A cognitive theory of juror decision making: The story model. *Cardozo Law Review, 13*, 519-558.

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.

Petty, R. E., & Cacioppo, J. T. (1979). Effects of forewarning of persuasive intent and involvement on cognitive responses and persuasion. *Personality and Social Psychology Bulletin, 5*, 173-176.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology, 41*, 847-855.

Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123-193.

Pierro, A., Mannetti, L., Erb, H. P., Spiegel, S., & Kruglanski, A. W. (2005). Informational length and order of presentation as determinants of persuasion. *Journal of Experimental Social Psychology, 41*, 458-469.

Platzer, C., & Bröder, A. (2012). Most people do not ignore salient invalid cues in memory-based decisions. *Psychonomic Bulletin & Review, 19*, 654-661.

Plott, C. R., & Smith, V. L. (Eds.). (2008). *Handbook of Experimental Economics Results* (Vol. 1). Amsterdam: Elsevier.

Pohl, R. F. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19*, 251-271.

Pohl, R. F. (2011). On the use of recognition in inferential decision making: An overview of the debate. *Judgment and Decision Making, 6*(5), 423-438.

Postmes, T., Spears, R., & Lea, M. (2002). Intergroup differentiation in computer-mediated communication: Effects of depersonalization. *Group Dynamics-Theory Research and Practice, 6*(1), 3-16.

Prentice, D. A., Gerrig, R. J., & Bailis, D. S. (1997). What readers bring to the processing of fictional texts. *Psychonomic Bulletin & Review, 4*(3), 416-420.

Rae, G. (1976). Table of A'. *Perceptual and Motor Skills*, *42*, 98.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59-108.

Ratcliff, R., Schmiedek, F., & McKoon, G. (2008). A diffusion model explanation of the worst performance rule for reaction tune and IQ. *Intelligence, 36*(1), 10-17.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition, 8*, 338-342.

Reinhard, M. A. (2010). Need for cognition and the process of lie detection. *Journal of Experimental Social Psychology, 46*(6), 961-971.

Reinhard, M. A., & Sporer, S. L. (2008). Verbal and nonverbal behaviour as a basis for credibility attribution: The impact of task involvement and cognitive capacity. *Journal of Experimental Social Psychology, 44*(3), 477-488.

Reinhard, M. A., & Sporer, S. L. (2010). Content versus source cue information as a basis for credibility judgments: The impact of task involvement. *Social Psychology, 41*(2), 93-104.

Reinhard, M. A., Sporer, S. L., & Marksteiner, T. (2008). Situational familiarity and the use of nonverbal and verbal information in judgments of veracity. *International Journal of Psychology, 43*(3-4), 617-617.

Richardson, D. C., Street, C. N. H., Tan, J., Kirkham, N. Z., Hoover, M. A., & Ghane Cavanaugh, A. (2012). Joint perception: Gaze and social context. *Frontiers in Human Neuroscience*.

Richter, T., Schroeder, S., & Wohrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology, 96*(3), 538-558.

Richter, T., & Späth, P. (2006). Recognition is used as one cue among others in judgment and decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 150-162.

Riek, B. M., Mania, E. W., & Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: A meta-analytic review. *Personality and Social Psychology Review, 10*(4), 336-353.

Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics and how can we tell? In G. Gigerenzer & P. M. Todd & The ABC Research Group (Eds.), *Simple Heuristics That Make us Smart* (pp. 141-167). New York: Oxford University Press.

Roberts, M. J., & Newton, E. J. (2001). Inspection times, the change task, and the rapid-response selection task. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 54*, 1031-1048.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review, 108*(2), 370-392.

Roggensack, K. E., & Sillars, A. (2013). Agreement and understanding about honesty and deception rules in romantic relationships. *Journal of Social and Personal Relationships*, 1-22.

Roggeveen, A. L., & Johar, G. V. (2002). Perceived source variability versus familiarity: Testing competing explanations for the truth effect.

Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review, 3*(1), 16-20.

Rouder, J., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225-237.

Sagarin, B. J., Rhoads, K. v. L., & Cialdini, R. B. (1998). Deceiver's distrust: Denigration as a consequence of undiscovered deception. *Personality and Social Psychology Bulletin, 24*(11), 1167-1176.

Saykaly, C., Talwar, V., Lindsay, R. C. L., Bala, N. C., & Lee, K. (2013). The influence of multiple interviews on the verbal markers of children's deception. *Law and Human Behavior, 37*(3), 187-196.

Schacter, D. L., & Tulving, E. (1994). What are the memory systems of 1994? In D. L. Schacter & E. Tulving (Eds.), *Memory Systems* (pp. 1-38). Cambridge, MA: MIT Press.

Scheepers, D., Spears, R., Doosje, B., & Manstead, A. S. R. (2006). The social functions of in-group bias: Creating, confirming, or changing social reality. *European Review of Social Psychology, 18*, 359-396.

Schroeder, S., Richter, T., & Hoever, I. (2008). Getting a picture that is both accurate and stable: Situation models and epistemic validation. *Journal of Memory and Language, 59*(3), 237-255.

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

Schul, Y. (1993). When warning succeeds: The effect of warning on success in ignoring invalid information. *Journal of Experimental Social Psychology, 29*(1), 42-62.

Schul, Y., Mayo, R., & Burnstein, E. (2004). Encoding under trust and distrust: The spontaneous activation of incongruent cognitions. *Journal of Personality and Social Psychology, 86*(5), 668-679.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61*, 195-202.

Schwarz, N., Sanna, L. J., Skurnik, I., & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology, 29*, 127-161.

Schwarz, N., Strack, F., Hilton, D. J., & Naderer, G. (1991). Base rates, representativeness, and the logic of conversation: The contextual relevance of "irrelevant" information. *Social Cognition, 9*, 67-84.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences, 10*(2), 70-76.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition, 88*, B11-B21.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies in self-reported lies. *Human Communication Research, 36*, 2-25.

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin, 137*, 207-222.

Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1954/1961). *The Robbers Cave Experiment: Intergroup Conflict and Cooperation*. Connecticut: Wesleyan University Press.

Sherman, J. W. (2006). On building a better process model: It's not only how many, but which ones and by which means. *Psychological Inquiry, 17*, 173-184.

Shiffrin, R. M., & Dumais, S. T. (1981). The development of automatism. In J. R. Anderson (Ed.), *Cognitive Skills and Their Acquisition* (pp. 111-140). Hillsdale, NJ: Erlbaum.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*, 1-19.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin, 105*, 131-142.

Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research, 31*(4), 713-724.

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

Sloman, S. A. (2002). Two systems of reasoning. In T. Gilovich & D. Griffin & D. Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (pp. 379-398). Cambridge: Cambridge University Press.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review, 4*(2), 108-131.

Smith, E. R., & Henry, S. (1996). An ingroup becomes part of the self: Response time evidence. *Personality and Social Psychology Bulletin, 22*, 635-642.

Snyder, M., & Campbell, B. (1980). Testing hypotheses about other people: The role of the hypothesis. *Personality and Social Psychology Bulletin, 6*(3), 421-426.

Snyder, M., & Swann, W. B. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology, 36*(11), 1202-1212.

Sperber, D., Clement, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*(4), 359-393.

Spinoza, B. (1677/1982). The Ethics and Selected Letters. In S. Feldman & S. Shirely (Eds.). Indianapolis, IN: Hackett.

Spivey, M., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences, 102*(29), 10393-10398.

Spivey, M. J. (2007). *The Continuity of Mind*. New York ; Oxford: Oxford University Press.

Spivey, M. J., & Dale, R. (2006). Continuous dynamics in real-time cognition. *Current Directions in Psychological Science, 15*(5), 207-211.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology, 45*(4), 447-481.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law, 7*(1), 36-97.

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 20*(4), 421-446.

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology Public Policy and Law, 13*(1), 1-34.

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

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology, 53*, 1-26.

Stiff, J. B., Kim, H. J., & Ramesh, C. N. (1992). Truth biases and aroused suspicion in relational deception. *Communication Research, 19*(3), 326-345.

Stiff, J. B., Miller, G. R., Sleight, C., Mongeau, P., Garlick, R., & Rogan, R. (1989). Explanations for visual cue primacy in judgments of honesty and deceit. *Journal of Personality and Social Psychology, 56*(4), 555-564.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 4*, 108-131.

Street, C. N. H., & Richardson, D. C. (in prep). Focal effect: Rethinking 'implicit lie detection'.

Street, C. N. H., Tbaily, L., Baron, S., Khalil-Marzouk, P., Hanby, B., Wright, K., & Richardson, D. C. (2011, April). Bloomsbury deception set, *British Psychological Society Division of Forensic Psychology Conference*. Portsmouth, UK.

Strömwall, L. A., & Granhag, P. A. (2005). Children's repeated lies and truths: Effects on adults' judgments and reality monitoring scores. *Psychiatry, Psychology and Law, 12*(2), 345-356.

Strömwall, L. A., & Granhag, P. A. (2007). Detecting deceit in pairs of children. *Journal of Applied Social Psychology, 37*(6), 1285-1304.

Strömwall, L. A., Granhag, P. A., & Hartwig, M. (2004). Practitioners' beliefs about deception. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception Detection in Forensic Contexts* (pp. 229-250). Cambridge: Cambridge University Press.

Strömwall, L. A., Granhag, P. A., & Jonsson, A. C. (2003). Deception among pairs: "Let's say we had lunch and hope they will swallow it!". *Psychology, Crime & Law, 9*(2), 109-124.

Swann, W. B., Guiliano, T., & Wegner, D. M. (1982). Where leading questions can lead: The power of conjecture in social interaction. *Journal of Personality and Social Psychology, 42*, 1025-1035.

Swann, W. B., Stein-Seroussi, A., & McNulty, S. E. (1992). Outcasts in a white-lie society: The enigmatic worlds of people with negative self-conceptions. *Journal of Personality and Social Psychology, 62*, 618-624.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*(5), 96-102.

Tajfel, H. (1981). *Human Groups and Social Categories: Studies in Social Psychology*. Cambridge: Cambridge University Press.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1*(2), 149-177.

Tajfel, H., & Turner, J. C. (1979). The social identity theory of intergroup behaviour. In W. G. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations* (pp. 7-24). Monterey, CA: Brooks/Cole.

The Global Deception Team (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37*(1), 60-74.

The Mathworks, Inc. (2000). MATLAB version 7.10.0. Natick, MA.

Thelen, E., & Smith, L. B. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

Thompson, V. A. (2013). Why it matters: The implications of autonomous processes for dual process theories--Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science, 8*(3), 253-256.

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

Todd, P. M., & Gigerenzer, G. (2005). Bounding rationality to the world. *Journal of Economic Psychology, 24*, 143-165.

Toris, C., & DePaulo, B. M. (1985). Effects of actual deception and suspiciousness of deception on interpersonal perceptions. *Journal of Personality and Social Psychology, 47*(5), 1063-1073.

Trabasso, T., Rollins, H., & Shaughnessy, E. (1971). Storage and verification stages in processing concepts. *Cognitive Psychology, 2*, 239-289.

Trope, Y., & Bassok, M. (1982). Confirmatory and diagnosing strategies in social information gathering. *Journal of Personality and Social Psychology, 43*(1), 22-34.

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

Ulatowska, J. (2013). Different questions - different accuracy? The accuracy of various indirect question types in deception detection. *Psychiatry, Psychology and Law*.

Unkelbach, C., Bayer, M., Alves, H., Koch, A. S., & Stahl, C. (2010). Fluency and positivity as possible causes of the truth effect. *Consciousness and Cognition, 20*(3), 594-602.

Uysal, A., & Oner-Ozkan, B. (2007). A self-presentational approach to transmission of good and bad news. *Social Behavior and Personality, 35*, 63-78.

van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias. *Psychological Science, 19*, 1131-1139.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.

van Ravenzwaaij, D., Mulder, M. J., Tuerlinckx, F., & Wagenmakers, E. J. (2012). Do the dynamics of prior information depend on task context? An analysis of optimal performance and an empirical test. *Frontiers in Psychology, 3*, 1-15.

van Swol, L. M., Malhotra, D., & Braun, M. T. (2012). Deception and its detection: Effects of monetary invesntives and personal communication history. *Communication Research, 39*(2), 217-238.

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2005). A dual-process specification of causal conditional reasoning. *Thinking & Reasoning, 11*, 239-278.

Voci, A. (2006). The link between identification and in-group favouritism: Effects of threat to social identity and trust-related emotions. *British Journal of Social Psychology, 45*, 265-284.

Vrij, A. (2008). *Detecting Lies and Deceit: Pitfalls and Opportunities* (2nd ed.). Hoboken, N.J.: Wiley ; Chichester : John Wiley.

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition, 1*(2), 110-117.

Vrij, A., Granhag, P. A., & Mann, S. (2010). Good liars. *The Journal of Psychiatry & Law, 38*(12), 77-98.

Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest, 11*(3), 89-121.

Wegner, D. M., Coulton, G. F., & Wenzlaff, R. (1985). The transparency of denial: Briefing in the debriefing paradigm. *Journal of Personality and Social Psychology, 49*(2), 338-346.

Weld, H. P., & Danzig, E. R. (1940). A study of the way in which a verdict is reached by a jury. *The American Journal of Psychology, 53*(4), 518-536.

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(6), 1235-1253.

Wickham, R. E. (2013). Perceived authenticity in romantic partners. *Journal of Experimental Social Psychology, 49*(5), 878-887.

Witte, R. S. (1964). Long-term effects of patterned reward schedules. *Journal of Experimental Psychology, 68*, 588-594.

Wood, W., Kallgren, C. A., & Preisler, R. M. (1985). Access to attitude-relevant information in memory as a determinant of persuasion: The role of message attributes. *Journal of Experimental Social Psychology, 21*, 73-85.

Wood, W., & Quinn, J. M. (2003). Forewarned and forearmed? Two meta-analytic syntheses of forewarnings of influence appeals. *Psychological Bulletin, 129*(1), 119-138.

Worchel, S., Rothgerber, H., Day, E. A., Hart, D., & Butemeyer, J. (1998). Social identity and individual productivity within groups. *British Journal of Social Psychology, 37*, 389-413.

Worth, L. T., & Mackie, D. M. (1987). The cognitive mediation of positive affect in persuasion. *Social Cognition, 5*, 76-94.

Wyer, R. S., & Frey, D. (1983). The effects of feedback about self and others on the recall and judgment of feedback-relevant information. *Journal of Experimental Social Psychology, 19*, 540-559.

Zellner, A., & Siow, A. (1980). Posterior odds ratios for selected regression hypotheses. In J. M. Bernardo & M. H. DeGroot & D. V. Lindley & A. F. M. Smith (Eds.), *Bayesian Statistics: Proceedings of the First International Metting held in Valencia (Spain)* (pp. 585-603). Valencia, Spain: University Press.

Zuckerman, M., DeFrank, R. S., Hall, J. A., Larrance, D. T., & Rosenthal, R. (1979). Facial and vocal cues of deception and honesty. *Journal of Experimental Social Psychology, 15*(4), 378-396.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 14, pp. 1-60). New York: Academic.

Zuckerman, M., Fischer, S. A., Osmun, R. W., Winkler, B. A., & Wolfson, L. R. (1987). Anchoring in lie detection revisited. *Journal of Nonverbal Behavior, 11*(1), 4-12.

Zuckerman, M., Kernis, M. R., Driver, R., & Koestner, R. (1984). Segmentation of behavior: Effects of actual deception and expected deception. *Journal of Personality and Social Psychology, 46*(5), 1173-1182.

Zuckerman, M., Koestner, R., Colella, M. J., & Alton, A. O. (1984). Anchoring in the detection of deception and leakage. *Journal of Personality and Social Psychology, 47*(2), 301-311.

Zukier, H., & Pepitone, A. (1984). Social roles and strategies in prediction: Some determinants of the use of base-rate information. *Journal of Personality and Social Psychology, 47*, 349-360.

Zwaan, R. A., & Radvansky, G. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162-185.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168-171.