EXPLICITLY AWARE OF CONFLICT:

CHALLENGING THE IMPLICIT CONFLICT DETECTION INTERPRETATION

OF THE BASE-RATE NEGLECT TASK

A Thesis Submitted to the College of Graduate and Postdoctoral Studies

In Partial Fulfillment of the Requirements

For the Degree of Master of Arts

In the Department of Psychology

University of Saskatchewan

Saskatoon

By

IAN NEWMAN

PERMISSION TO USE

DISCLAIMER

ABSTRACT

Conflict reasoning problems cue two competing responses to the problem, requiring the reasoner to resolve the conflict; non-conflict problems cue the same response. The central claim of the conflict detection literature is that conflict is detected implicitly without explicit awareness. The goal of this research is to test the hypothesis that reasoners are explicitly aware of the conflict with the base-rate reasoning task. Base-rate neglect is the tendency to undervalue base-rate ratios in favour of stereotypical personality descriptions. Conflict is studied with the base-rate task by pitting probabilistic information (the ratio) against believable information (the stereotype); performance is measured on conflict problems relative to non-conflict problems. In this research, the extremity of the base-rate ratios was manipulated and a neutral problem condition was included. Behavioural measures of confidence ratings, response times, and eye-gaze fixation times were collected. Retrospective self-reports were taken regarding awareness of conflict in the problems and conflict resolution strategy. In two experiments, there was compelling evidence that reasoners are more explicitly aware of conflict than previously assumed, that base-rate neglect is a function of conflict resolution strategy, and that the presumed indices of conflict detection index more than detection, namely, processes of conflict resolution and recognition of coherent information. This evidence provides a strong challenge to the predominant conflict-detection interpretation.

ACKNOWLEDGEMENTS

I would like to thank the members of my advisory committee, Dr. Jamie Campbell and Dr. Steve Prime for their assistance, direction and insightful feedback, and my external examiner Dr. Tim Claypool, for his time dedicated and feedback contributed to this thesis. I want to thank the Natural Sciences and Engineering Research Council and the College of Graduate and Postdoctoral Studies for funding this project. A special thanks to Prachi, Carly King, Hina Mahboob, and Jesse Ponath for their assistance with data collection. Lastly, I would like to extend my eternal thanks to Dr. Valerie Thompson for her expert supervision, guidance, and support with this project.

DEDICATION

I dedicate this thesis to my daughter, Kaia, the inspiration and motivation for all my present and future achievements.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction

Many situations cue multiple, competing responses to a problem. The goal of conflict detection research is to understand what cues a reasoner to avoid biased thinking and engage in analytic thinking. It is believed that the primary cues for analytic thought are detection of conflict (De Neys, 2014; Pennycook, Fugelsang, & Koehler, 2015) and affective metacognitive states (Ackerman & Thompson, 2017; Thompson, Prowse-Turner, & Pennycook, 2011; Thompson et al., 2013). The key questions for researchers are whether reasoners detect the fact that there are competing responses and whether reasoners are aware of the conflict between said responses. Current theories stress that detection of competing responses is a cognitively implicit process (De Neys, 2014; Handley & Trippas, 2015; Pennycook et al., 2015) experienced only as an affective "gut-feeling" without any explicit awareness (De Neys, 2014).

The purpose of this research is to test the central claims of the *conflict-detection* interpretation (De Neys, 2012; 2014): reasoners implicitly detect but are explicitly unaware of conflicting responses and that the indices of conflict detection (e.g., confidence reports, response times, eye-gaze fixation times) index detection *per se*. The alternative interpretation proposed here is that reasoners are explicitly aware of the conflict and that the indices of conflict detection index more than detection, namely, cognitive factors of conflict resolution and monitoring of information *coherence* (the consistency of the information, assessed intuitively; Topolinski, 2011), hereafter referred to as the *coherence-resolution* interpretation. The coherence-resolution interpretation provides a strong challenge to the predominant conflict-detection interpretation.

1.1. Base-Rate Neglect

Conflict reasoning problems present reasoners with two inconsistent sources of information intended to cue two competing responses. These problems typically pit the rules of logic and/or probability against the person's *a priori* beliefs, where the normatively correct response is to apply logical or probabilistic rules. As a control, in non-conflict problems both pieces of information cue the same response.

The *base-rate neglect* phenomenon is the tendency to undervalue statistical information in favour of a personality description when estimating the probability of category membership (Kahneman & Tversky, 1973; Tversky & Kahneman, 1974). The base-rate neglect reasoning task is frequently used to study conflict detection. The task involves estimating the probability of category membership based on the information provided in each problem: the category base-rate ratio and a stereotypical personality description, e.g.,

> In a study, 1000 people were tested. Jack is a randomly chosen participant of this study. Among the participants there were 5 engineers and 995 lawyers. Jack is 36 years old. He is not married and is somewhat introverted. He likes to spend his free time reading science fiction and writing computer programs. What is the probability that Jack is an engineer?

The reasoner is asked to estimate the likelihood that an individual (e.g., Jack) belongs to one of the two categories (e.g., engineer) given the probabilistic information and the stereotypical personality description (e.g., this person in question closely matches the stereotype of an engineer). The probability and description can suggest two competing responses (conflict, as above) or one consistent response (non-conflict).

Probability estimates for conflict problems are typically further from the base-rate probability than responses for non-conflict problems (Kahneman & Tversky, 1973; De Neys & Glumicic, 2008; Tversky & Kahneman, 1974). Presumably, this occurs because people tend to focus on the personality descriptions and ignore or minimize the base-rate ratios on conflict problems (Barbey & Sloman, 2007). Conflict is removed from the problems by reversing the base-rate ratio values (e.g., 5 engineers and 995 lawyers is changed to 995 engineers and 5 lawyers).

1.2. Conflict Detection

Reasoners consistently appear to neglect the base-rate ratios for conflict problems when estimating probabilities but there is evidence that they do not completely ignore the ratios. Conflict detection studies attempt to determine if reasoners are sensitive to the conflict between competing responses that are inconsistent with one another, such as the stereotype and base-rate responses in the above example (De Neys, 2014). Sensitivity to conflict is operationalized as performance on conflict problems relative to non-conflict problems. Even when a reasoner's probability estimates suggest that the base-rate was neglected, other elements of their performance indicate that they were, on some level, sensitive to the conflict between the base-rate ratio and the stereotype.

There is considerable evidence that reasoners display myriad behavioural indicators of sensitivity to conflict even when their responses do not (see De Neys, 2014 for a review). Relative to non-conflict problems, on conflict problems, reasoners typically display longer response times (Bonner & Newell, 2010; De Neys & Glumicic, 2008; Thompson et al., 2011), lower confidence (Pennycook, Trippas, Handley, & Thompson, 2014; Thompson & Johnson, 2014; Thompson et al., 2011), increased autonomic arousal (De Neys, Moyens, & Vansteenwegen, 2010), and increased visual attention on the relevant, conflicting problem information (Ball, Phillips, Wade, & Quayle, 2006). Moreover, this has been found on a variety of reasoning tasks: base-rate neglect (De Neys & Glumicic, 2008), conjunction fallacy (De Neys, Cromheeke, & Osman, 2011), ratio-bias (Mevel et al., 2015), syllogistic reasoning (Thompson &

Johnson, 2014), and the cognitive reflection test (De Neys & Feremans, 2012). Reasoners also display better recall (on surprise recall tasks) for base-rate ratios from conflict problems than from non-conflict problems (De Neys & Glumicic, 2008; Franssens & De Neys, 2009), suggesting that reasoners notice the base-rate ratios that conflict with the personality descriptions.

As an index of conflict detection, visual attention has been relatively underutilized in reasoning research. Using a belief-bias task, Ball, Philips, Wade, and Quayle (2006) found that inspection time increased on problem premises for conflict problems relative to non-conflict problems, suggesting that the response to conflict is to review the conflicting information. De Neys and Glumicic (2008) applied a paradigm where base-rate ratios were visible while a button was pressed but hidden otherwise. In their task, reasoners reviewed the base-rate ratios more for conflict trials than non-conflict trials. To date, conventional eye-tracking measures have not been applied to a base-rate reasoning task.

The crucial evidence here is that these behavioural measures are found independent of performance on the task itself: reasoners that give responses consistent with the stereotype nonetheless display these conflict detection effects (Bonner & Newell, 2010; De Neys, 2012; Morsanyi & Handley, 2012). The prevailing conflict-detection interpretation is that these measures index conflict detection, detection is an implicit process (De Neys, 2012; Pennycook et al., 2015), and reasoners are unaware of the conflict at an explicit level (De Neys, 2014). De Neys (2014) contends that conflict gives rise to arousal, which in turn is noticed by reasoners; the nature of the conflict is not explicitly understood beyond a “gut-feeling” (De Neys et al., 2010). The affective experience is merely a signal to question one’s initial response (De Neys, 2012, Thompson et al., 2011).

The conflict-detection interpretation rests on the assumption that if a reasoner gives the normatively incorrect response to a problem, they must have been unaware of the conflict. Otherwise, they would have given the statistically correct response. Thus, they must have processed the conflicting information at an implicit level (eliciting the aforementioned behavioural responses) but because they gave the incorrect response, the conflict was not processed at an explicit level of awareness. The source of the conflict is thought not to be accessible to the reasoner. To date, this assumption has not been tested.

1.3. Conflict Resolution

The conflict-detection interpretation relies on the assumption that when one gives a non-normative response to a reasoning problem, it is indicative of a reasoning failure (De Neys, 2014, Pennycook et al., 2015). It is assumed that when a response other than the normatively correct one is given, it results either from (a) a failure to detect the conflict or (b) successful implicit detection but a failure to calculate the correct response (from the above example, 0.5% chance that Jack is an engineer; Pennycook et al., 2015). As argued below, it is not necessary to assume that conflict detection is implicit.

Alternatively, it could be argued that an incorrect response is indicative of a reasoning strategy (Pennycook & Thompson, 2012), intention (e.g., rationalization of the statistically improbable response; Pennycook et al., 2015), or a function of relative weighting of information. In other words, reasoners could deliberately choose the personality description over the base-rate ratio. For example, on a base-rate problem, a reasoner might disregard the probability because he felt his personal experience with the information provided by the stereotype was more reliable to reach a solution. In this case, one might well be aware of the probabilistic answer and aware of the conflict but ultimately select a non-normative response. There is no need to assume that conflict detection is implicit nor that an incorrect answer indicates the absence of awareness. The nature of the reasoning process may hinge less on the implicit detection of conflict and more on the explicit resolution of conflict based on subjective weighting of information in the problem. Thus, reasoners would be generally aware of conflict at an explicit level.

Pennycook and Thompson (2012) argued that reasoners strategically choose which source of information on which to base their responses. Their interpretation was that the base-rate neglect phenomenon arises when the combination of two strategies (reliance on ratios or reliance on descriptions) are averaged in group level analyses. Under this interpretation, it is argued that base-rate neglect occurs because generally, those who apply a strategy of reliance on the personality description are more numerous (Newman, Gibb, & Thompson, 2017; Pennycook & Thompson, 2012). This interpretation is consistent with an extensive line of research that indicates there is a significant minority of reasoners who demonstrate frequently successful normative reasoning: giving responses consistent with the rules of logic and probability (Evans & Stanovich, 2013; Stanovich & West, 2000; Toplak, West, & Stanovich, 2011; Thompson & Johnson, 2014). Essentially, base-rate neglect can be thought of as the proportion of conflict resolution strategies applied by reasoners; the majority demonstrate the phenomenon but a minority primarily do not.

1.4. Indices of Conflict Detection, Coherence, and Resolution

Conflict effects can be explained without the assumption of implicit conflict detection. There are several potential reasons why conflict problems elicit longer response times and lower confidence (and others; De Neys, 2014). A reasoner that is aware of conflict must choose between two competing responses. Merely selecting between two alternatives may increase time to respond compared to non-conflict problems that cue only one response. Moreover, awareness that there is more than one plausible response to the problem forces the reasoner to choose between them; any indecision or uncertainty regarding which to select could reduce confidence. The reasoner may second guess their choice and then rate their confidence as lower.

There is reason to question whether the indices of conflict detection singularly index the implicit detection of conflict. Measures such as self-reported confidence and response times may also index the resolution of conflict or monitoring of information coherence. If these measures are sensitive to multiple cognitive factors, inferences regarding implicit detection of conflict are unwarranted. For example, if reasoners display lower confidence on conflict problems, the source of the reduction in confidence is unclear; it could be caused by conflict detection, recognition of incoherence of information, the resolution of conflict, or a combination of these factors. Similarly, one cannot infer that observable conflict effects across multiple behavioural measures indicate that these measures all reflect the same cognitive conflict detection process. It is assumed that conflict effects on these behavioural measures reflect the same underlying cognitive processes (De Neys, 2014); as argued below, that assumption may also be unwarranted.

1.5. Explicit Awareness of Conflict

Reasoners may be more explicitly aware of conflict than previously believed. To test that hypothesis, retrospective self-report measures of conflict awareness and conflict resolution strategy provide the first attempt to create indices of explicit conflict awareness that can be used alongside the presumed indices of conflict detection. Under the conflict-detection interpretation, implicit and explicit measures should not necessarily correlate with one another because detection is believed to be processed implicitly; the reasoner is not explicitly aware of the conflict. Thus, how a reasoner reports their own explicit awareness would not necessarily be associated with their behaviour that indicates conflict was implicitly detected.

Alternatively, the coherence-resolution interpretation would predict that implicit and explicit measures will correlate with each other because the presumed indices of conflict detection are also measures of conflict resolution processes. Under this interpretation, reasoners are more explicitly aware than previously believed; therefore, reports of awareness will be related to behavioural reactions to conflict. The rationale is that these measures are all indexing the same cognitive construct, namely, relative awareness of conflict.

In the present research, self-reports were used to assess explicit awareness of conflict. The validity of self-report data in problem-solving research has been challenged; the reports are frequently inaccurate because they require introspective inferences about covert cognitive processes (Cohen, 1987; Ericsson & Simon, 1987). In that case, what are the conditions that allow for reliable self-reports of cognitive processes (Taylor & Dionne, 2000)? Self-reports access the content of consciousness, but not the cognitive processes that never reach consciousness (Wilson, 1994). Also, there is a general consensus that retrospective protocols are subject to some forgetting (Russo, Johnson, & Stephens, 1989). However, this has been found to be a greater concern for low-level cognitive processes (e.g., basic perceptual processing) than higher-order processes (e.g., strategy use; Russo et al., 1989). Questionnaires should focus on "what" rather than "why" queries, a matter-of-fact approach to the questions (instead of evaluative), and opportunities for clarification from the respondents; these conditions allow for collection of valid and reliable retrospective report data (Taylor & Dionne, 2000). These prescriptions were followed here to maximize the validity of the self-report measures: matter-of-fact questions regarding higher-order cognitive processes with opportunity for clarification, wherever possible.

Retrospective self-report protocols appear to be an adequate method of assessing higher-order cognitive processes. Two measures of explicit awareness were gathered: the strategy adopted to resolve conflict (relying on the probabilistic information or relying on the believable information) and the explicit awareness of conflict. One option was to query reasoners on strategy and awareness on a per-problem basis. The concern was the reactivity of the questions, alerting the reasoner of the possibility of conflict in the coming problems and potentially influencing their performance on subsequent trials during the task (e.g., increased response times; Russo et al., 1989). Therefore, a retrospective self-report at the end of the experiment was the most appropriate measure to test the hypotheses regarding explicit awareness of conflict and conflict resolution strategies. Additionally, strategy reports can be verified by the behavioural data: the behavioural responses should map sensibly onto the reported strategies applied by reasoners.

1.6. Objectives

The main hypotheses of the present research were (a) that reasoners are explicitly aware of conflict in the reasoning problems they solve, (b) the presumed indices of conflict detection are sensitive to other cognitive processes that elicit similar behavioural responses as conflict detection, and (c) base-rate neglect is explained by individual differences in conflict resolution strategies. To test these hypotheses, two experiments were conducted using the base-rate task. In Experiment 1, novel self-report measures were used to assess explicit awareness of conflict; these measures have not been used in conflict detection research. A novel, retrospective self-report was also used to determine reasoning strategy and was corroborated with behavioural measures of probability estimates, confidence reports, response times, and eye-gaze fixation times. Eye-tracking measures have not been used in base-rate neglect research, to date. Another version of non-conflict problems (i.e., neutral problems, see below) were included to test the sensitivity of the behavioural indices to cognitive factors unrelated to conflict detection. In Experiment 2, the explicit awareness and strategy self-reports were again used alongside two attempts to experimentally manipulate explicit awareness. The first was an intervention questionnaire aimed at increasing awareness by alerting reasoners to the conflict in the problems. The second was a concurrent spatial memory task intended to load working-memory and reduce the ability to explicitly (but not implicitly) detect the conflict in the problems.

Chapter 2. Experiment 1

Reasoners performed a base-rate task using the two-response paradigm (see below; Thompson et al., 2011); the format of this task closely matched the base-rate task used by Newman et al., (2017). Probability estimates, response times, and confidence ratings were collected for each problem. Additionally, eye-gaze fixation times were measured for two areas of interest (AOI) in the problem text presented to reasoners: the area surrounding the base-rate ratio and the area surrounding the personality description. Retrospective self-reports were also taken at the end of the task, assessing explicit awareness of conflict and the strategy adopted to solve the preceding problems.

2.1. Neutral Personality Descriptions

Congruency between description and base-rate ratio was manipulated. The description and base-rate suggested one consistent response (congruent), competing, inconsistent responses (incongruent), or contained an uninformative personality description (neutral). De Neys and Glumicic (2008) altered personality descriptions in their base-rate task to include neutral control problems; for these problems, the personality descriptions offered uninformative details (e.g., height, eye-colour, city of residence instead of a stereotype) regarding category membership. Probability estimates on neutral base-rate problems have been found to be further from the base-rate probability than non-conflict problems yet closer to the base-rate probability than conflict problems (De Neys & Glumicic, 2008; Newman et al., 2017; Pennycook & Thompson, 2012; Thompson et al., 2011).

Neutral problems do not contain conflicting information. If the difference in behavioural responses (e.g., lower confidence, longer response times, longer fixation times) between conflict and non-conflict problems were genuine indicators of conflict detection, then the effect of conflict would be more pronounced for conflict problems than for neutral problems. Alternatively, previous research has found that neutral problems are associated with lower confidence ratings than both congruent and incongruent problems that contain stereotypical personality descriptions (Thompson et al., 2011). This evidence suggests that confidence is indexing other cognitive processes than conflict detection. For example, because the personality descriptions in the neutral problems do not cue a believable response (De Neys & Glumicic, 2008), reasoners may view the information in neutral problems as less coherent than congruent and incongruent problems because the description is not consistent with either category in the base-rate ratio. Coherence is thought to elicit fluently generated responses and, in turn, higher confidence (Koriat, 2012; Thompson et al., 2011). Thus, one explanation for low confidence for neutral problems is that the information is less coherent. Additionally, neutral problems could evoke feelings of uncertainty, as the descriptions do not offer useful information to solve the problem, a potential source of reduced confidence and longer response times. Therefore, if the presumed behavioural indicators of conflict detection are more pronounced for neutral problems than for incongruent problems (Thompson et al., 2011), it indicates that there are cognitive factors, other than conflict detection, that elicit these responses in the base-rate task.

2.2. Base-Rate Ratio Extremity

The extremity of the base-rate ratios was also manipulated, using extreme (995:5), moderate (700:300), and balanced (510:490) ratios to vary the degree of conflict (Newman et al., 2017; Pennycook, Fugelsang, & Koehler, 2012; Pennycook et al., 2015). Newman, Gibb, and Thompson (2017), using the two-response paradigm (Thompson et al., 2011) found that probability estimates varied between base-rate ratio extremity conditions for neutral and non-conflict problems, reflecting an awareness of differences in the base-rate probability. On conflict trials, this difference only emerged for extreme ratios. Pennycook, Fugelsang, and Koehler (2015) also found that more responses consistent with the base-rate probability were given when base-rates were extreme compared to moderate.

Varying the extremity of the base-rate ratios effectively varies the degree of conflict in the problems. Conflict effects on confidence are typically found using extreme base-rates (De Neys & Glumicic, 2008; Pennycook & Thompson, 2012; Pennycook et al., 2014; Thompson et al., 2011). The extremity manipulation also varies the coherence of the information in the problems. In congruent, non-conflict problems, the personality description is more coherent with an extreme base-rate ratio than with a moderate or balanced ratio. Thus, the difference between conflict and non-conflict problem confidence could reflect increased confidence when information is more coherent instead of decreased confidence when conflict is detected (or a combination of both). If confidence is an index of conflict detection, then varying the base-rate ratio extremity would influence confidence on the incongruent conflict problems, but not on the non-conflict congruent and neutral problems. On the other hand, if confidence indexes coherence as well as conflict, then the effect of extremity would be detectable for congruent problems, as well.

2.3. Visual Attention

Eye-tracking has not been used to measure visual attention on the base-rate task. Ball et al. (2006) found that conflict problems in a syllogistic reasoning task elicited more fixation time on the problem premises than non-conflict problems. Under the conflict-detection interpretation, on conflict problems in the base-rate task, fixation times should increase (presumably to review the conflicting base-rate ratios and personality descriptions). Alternatively, the inspection time of the uninformative personality descriptions could be the longest, possibly driven by uncertainty or attempts to resolve the inconsistency.

Visual attention duration, response times, and confidence have all been presented as evidence for implicit conflict detection (De Neys, 2014), suggesting they all index the same underlying cognitive process. Therefore, the conflict-detection interpretation would predict that on conflict problems, response times and fixation times would increase (presumably to review the conflicting base-rate ratios and personality descriptions) and confidence would decrease, relative to non-conflict problems. Alternatively, if the response times, fixation times, and confidence patterns are not consistent with each other (e.g., response times are longest for conflict problems but confidence is lowest for neutral problems, not conflict problems) it would suggest that these measures index different cognitive factors. For instance, confidence could be sensitive to the coherence of the information and response and fixation times sensitive to the resolution strategies adopted by the reasoners.

Furthermore, if response times and visual attention index the resolution of conflict, then the reviewing of information on conflict trials should be a function of the resolution strategy adopted by the reasoner. Reasoners who favour probabilistic information would attend more to the base-rate ratios and less to the personality descriptions compared to those who favour believable personality descriptions, suggesting that measures of response time and fixation time index more than just a general response to implicitly detecting conflict, but also index conflict resolution processes.

2.4. Hypotheses

2.4.1. Base-rate neglect: probability estimates.

The probability estimate measure serves to demonstrate the base-rate neglect phenomenon, which was expected to be replicated: probability estimates would be further from the base-rate probability for incongruent problems than congruent problems. Probability estimates for neutral problems were also expected to replicate previous research (De Neys & Glumicic, 2008; Newman et al., 2017; Pennycook & Thompson, 2012; Thompson et al., 2011): further from the base-rate probability than estimates for congruent problems but closer to the base-rate probability than estimates for incongruent problems. The extremity effect was also expected to be replicated (Newman et al., 2017; Pennycook et al., 2015). The effect of base-rate extremity was expected to manifest across all extremity conditions for neutral and congruent problems, but only in the extreme condition for incongruent problems.

2.4.2. Hypothesis A: Behavioural indices are sensitive to multiple cognitive factors.

2.4.2.1. Confidence.

The degree of conflict was manipulated in two ways: congruency and extremity. The typical effect of conflict on confidence was expected to be replicated, where incongruent problems show lower confidence ratings than congruent problems. Under the conflict-detection interpretation, confidence would be lower for incongruent (conflict) problems than for neutral problems. On the other hand, the coherence-resolution interpretation predicts that confidence for neutral problems would be lower than for incongruent problems, suggesting that confidence is indexing more than conflict detection, such as the relative coherence of the information presented.

The manipulation of base-rate extremity also varies the conflict in the problems. In conflict problems, more extreme ratios present more conflict with the personality descriptions. For example, despite fitting the stereotypical description of an engineer, the probability Jack is an engineer is only 0.5% in extreme base-rate ratios but 30% in moderate base-rate ratios. Thus, conflict between the personality description and base-rate ratio is highest for incongruent-extreme problems; coherence between the personality description and base-rate ratio is highest for congruent-extreme problems. Under the conflict-detection interpretation, the effect of conflict would be largest for the incongruent-extreme problems. Conversely, if confidence is an index of coherence, the effect of extremity would also manifest for congruent problems. Not only would the decrease in confidence be greatest when conflict is highest (i.e., incongruent-extreme problems) but the increase in confidence would be greatest when coherence is highest (i.e., congruent-extreme).

2.4.2.2. Response times.

Under the conflict-detection interpretation, response time measures would be expected to mirror confidence measures: where confidence is decreased, response time is increased. Both measures have been argued to index conflict detection, reflecting the same underlying cognitive process. Therefore, response times for incongruent problems would be expected to be longer than response times for neutral problems; response times for congruent problems would be the shortest.

Under the coherence-resolution interpretation, the underlying cognitive processes indexed by confidence and response times may differ. Conflict resolution, uncertainty, coherence, and weighing comparative strength of information may impact response times (and other similar measures) differently. The prediction of the coherence-resolution interpretation is that the presence of neutral personality descriptions would evoke longer response times than for incongruent and congruent problems. The rationale is analogous to the confidence measure. The neutral descriptions provide little information upon which to base a decision, potentially generating feelings of uncertainty in reasoners and a more challenging conflict resolution decision, both of which would serve to increase response times. Moreover, coherent information is thought to elicit fast, fluent processing (Koriat, 2012); the less coherent information in the neutral problems reduces fluency, presumably also increasing response times.

2.4.2.3. Fixation times.

The conflict-detection interpretation would suggest that fixation times on both base-rates and personality descriptions would be longest for incongruent problems (converging with the response time data). The problems with the highest degree of conflict are where the cue to reinspect the problem would be strongest. On the other hand, if either response times or fixation times are longer for neutral problems than incongruent problems, it provides a challenge to the conflict-detection interpretation. Fixation time as an index of conflict detection should be most sensitive when conflict is present. Otherwise, fixation time indexes more than conflict detection and is not a trustworthy indicator of detection of conflict.

2.4.3. Hypothesis B: Reasoners are explicitly aware of conflict.

Retrospective ratings of explicit awareness of conflict were a novel measure constructed for this experiment and have not been used in reasoning research to date. Three questions were created. The first was a rating of awareness that there were two (or more) different answers available to the problems. The second asked if the reasoner noticed the numbers and personality sketches suggested different responses. The third posed the question of whether the reasoner ever felt the need to decide between two choices to select the best solution. Responses were rated on a seven-point Likert scale.

The expectation was that the three retrospective questions would correlate with one another, indicating that they assess the same underlying construct: explicit awareness of conflict. Furthermore, the *a priori* hypothesis was that reasoners are more explicitly aware of conflict than previously acknowledged by the conflict-detection interpretation. Thus, it was predicted that reasoners would report relatively high awareness of conflict and that these measures would correlate with the hallmark behavioural indicators of conflict detection. The correlation was expected because the behavioural indices and the self-report questions were predicted to both index detection of conflict, only differing on the dimension of conscious awareness. In summary, the conflict-detection interpretation expects these measures would not correlate (detection does not reach awareness) and the coherence-resolution interpretation expects these measures would correlate (both index relative awareness of conflict).

2.4.4. Hypothesis C: Base-rate neglect is a function of conflict resolution strategy.

Strategy self-reports were expected to discriminate participants based on their approach to conflict resolution. The question asked reasoners to report on their general approach to solve the problems: rely on the base-rate ratios, rely on the personality descriptions, or take a case-by-case approach (open ended options were also available). This question was a single retrospective query and a novel measure.

If the strategy report measure is a valid index of resolution strategy, then the effect of strategy would be evident across all the dependent measures in sensible patterns. Preliminary predictions included (a) the group that reports relying on the base-rate ratios would give responses closest to the statistically correct base-rate probability, (b) the group that reports relying on the personality descriptions would be less confident with uninformative neutral descriptions, and (c) the group that reports taking a case-by-case approach would take the longest to respond and report the highest explicit awareness of conflict, as this group was expected to consider the base-rate ratios and personality descriptions more equally than the other two groups. Visual attention was also expected to be a function of strategy: the group that reports relying on the base-rate ratios would attend less to the personality descriptions than the other groups while the group that reports relying on the personality descriptions would attend less to the base-rate ratios than the other groups. The coherence-resolution interpretation would predict that these strategies map sensibly on to the behavioural measures collected, indicating that that this retrospective question is a valid measure of resolution strategy and that (at least part of) what these measures index is the process of conflict resolution adopted by the reasoner.

2.5. Method

2.5.1. Participants.

One hundred twenty participants (74 female, 46 male, M_{age}=20.68 years) from the University of Saskatchewan took part in the study for partial course credit.

2.5.2. Apparatus and stimuli.

The task was performed on a Microsoft Windows laptop computer with a 1600x900 resolution display using the E-Prime Psychology Software Tools program (Psychology Software Tools, Pittsburgh, PA). Base-rate problems were presented as bitmap images of text on a black background. The problems described the ratio of membership for two categories (e.g., 995 engineers, 5 lawyers) within a sample of 1000 people and a stereotypical personality description of one individual who had been randomly selected from the sample (e.g., Jack is somewhat introverted and spends his free time writing computer programs and reading science fiction). The base-rate ratio was always presented above the personality description. Incongruent and congruent problems featured the same personality descriptions. Neutral problems contained descriptions that were uninformative regarding category membership (e.g., description includes age, height, and eye colour; modified from De Neys & Glumicic, 2008).

Nine problems from each congruency condition were solved by each participant. Three base-rate ratio extremities were also included: extreme (995:5), moderate (700:300), and balanced (510:490). Participants solved nine problems from each base-rate extremity condition and values within ratios had three variations (e.g., 995:5, 996:4, 997:3). Problem content was not repeated throughout the task for any participants. Congruency and extremity were counterbalanced across participants. The nine neutral problems were tested equally often in each base-rate extremity condition and were always neutral congruency. The remaining 18 problems were tested equally often as congruent and incongruent; these were also tested equally often in each of the three base-rate extremity conditions. For a full list of the base-rate problems, refer to Appendix A. To facilitate the use of an eye-tracker, the text of the base-rate problems was adjusted to isolate the base-rate ratios and personality descriptions from the rest of the text:

In a study, 1000 people were tested. Among the participants there were:

5 engineers

995 lawyers

Jack is a randomly chosen participant of this study.

Jack is 36 years old.

He is not married and is somewhat introverted.

He likes to spend his free time reading science fiction

and writing computer programs.

What is the probability that Jack is a lawyer?

Self-report measures were taken for the general strategy applied to solve the problems (the full list of self-report questions is in Appendix B). Five options were available:

1. I mostly relied on the numbers
2. I mostly paid attention to the personality sketches
3. I guessed randomly
4. I evaluated each problem on a case-by-case basis
5. None of the above.

Each response allowed for additional text to be provided by the participant to elaborate on their self-reported strategy. Responses were also gathered for three questions meant to index relative awareness of conflict:

1. Did you feel that there was two (or more) answers, which were quite different from one another, to some of the problems (Q1)?
2. Did you notice that sometimes the personality sketch and the numbers suggested different responses (Q2)?

3. Did you ever feel torn between two choices and needed to decide which you felt was a better solution (Q3)?

These responses were entered on a seven-point Likert scale (1 – Never, 7 – Always).

Participants were seated approximately 55 to 65cm from the laptop computer screen on a stationary chair. No chin rest was used. The SMI RED-m remote eye-tracker (SensoMotoric Instruments; http://www.smivision.com) was used that allows for user mobility and minimizes intrusiveness (Mele & Federici, 2012). The eye-tracker collected information regarding the frequency and order of eye-gaze fixations, frequency of gaze revisits to each area of interest, duration of each fixation, and pupil diameter. The threshold for a fixation was 80ms of sustained gaze (or longer) on a single position on the screen with a fixation radius within 1° of dispersion (a dispersion threshold found to produce reliable, replicable results; Blignaut, 2009). This dispersion criterion corresponded to approximately 0.96 to 1.13cm on the computer screen. A conventional fixation threshold used in eye-tracking research that has been found to effectively discriminate fixations from other oculomotor activity (Manor & Gordon, 2003) and that has been used previously in reasoning research (Ball et al., 2006) is 100ms. For the SMI BeGaze eye-tracking software, the number of false fixations becomes large when the threshold is under 40ms but the distribution of fixation duration and fixation frequency is very similar for any threshold 80ms and above (Holmqvist et al., 2011). The SMI BeGaze software also uses 80ms as the default fixation threshold. For these reasons, a fixation threshold of 80ms was considered acceptable for the purposes of this study.

2.5.3. Procedure.

Participants began the task with calibration of the eye-tracker, fixating their eye-gaze upon a target black circle on a grey background. The circle would disappear and reappear elsewhere on the screen once the eye-tracker recognized that the participant's eye-gaze was fixated upon the target. Calibration success was achieved when accuracy for x and y coordinates of the left and right eyes was within the accepted range of error: 1° of horizontal and vertical deviation from the target for each eye. On calibration failure, up to five retries were available. If all five calibration retries failed, the task would proceed without collection of eye-gaze data.

Base-rate problems were presented in a random order using the two-response paradigm (Thompson et al., 2011). Participants provided probability estimates (on a scale from 0 to 100); they were instructed to respond intuitively with the first answer that came to mind for their initial response to each trial (Time 1) and to take their time and think carefully for the final response to each trial (Time 2). Time 1 responses were also subject to time-pressure: problem text was white in colour and shifted to red after 11 seconds[1]. The colour change served as a prompt to enter a response within 1 second; participants were instructed to give their response immediately after the shift in text colour occurred. The purpose of this time pressure was to prevent participants from spending excessive amounts of time thinking about their Time 1 responses. For each probability estimate entered, participants rated their confidence in the response on a nine-point Likert scale (1 – Guessing, 9 – Certain I'm Right). Response times were also recorded for each probability estimate. Probability estimates were recorded once the participant entered the value between 0 and 100 and pressed the Enter key. The response time was recorded once the Enter key was pressed.

The trial progression is displayed in Figure 2.1. Each trial began with a central fixation cross presented for two seconds, followed by the base-rate problem for the Time 1 probability

[1] Problems remained on screen until a response was entered, regardless of whether the deadline was met. The deadline duration was previously determined to be challenging but afforded enough time for participants to read the problems in full (Thompson, Prowse-Turner, & Pennycook, 2011).

estimate. After 11 seconds, if a probability estimate had not been entered, the text colour changed from white to red. The problem in red text remained on screen until a response was provided. After the Time 1 probability estimate was entered, the Time 1 confidence rating was recorded. Next, the same base-rate problem was displayed for the Time 2 probability estimate to be recorded in free time; confidence in the Time 2 response was also rated by participants after the Time 2 probability estimate.

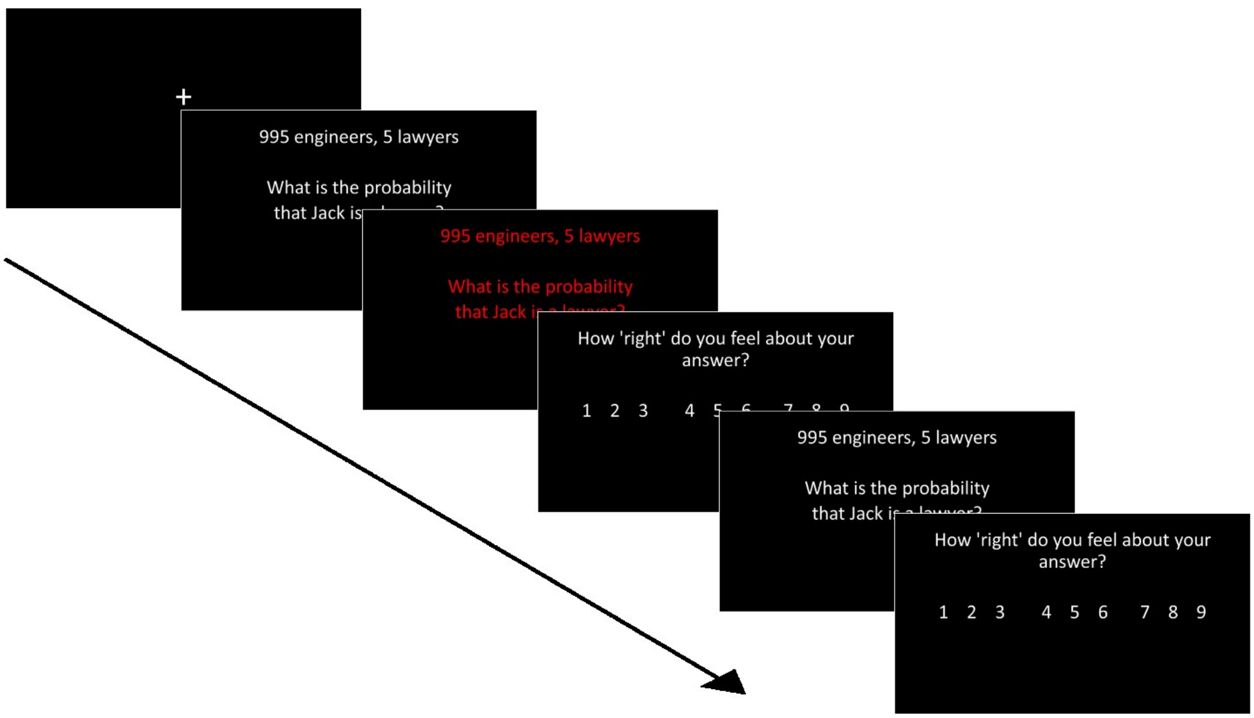

Figure 2.1. Event sequence for Experiment 1 trials (with simplified text for legibility; see pages 29-30 and Appendix A for format of the problem text).

Upon completion of all 27 base-rate problems, the explicit awareness questions and strategy assessment were taken. The strategy assessment was presented first. Subsequently, the explicit awareness questions were presented. These problems were presented sequentially for all participants in a block with filler questions (e.g., recognition memory questions). Completion of these self-report questions concluded the task. The task instructions are in Appendix C.

2.6. Results

The mean response time for Time 1 responses was 12.26 seconds. Three participants had mean Time 1 response times that were greater than 2.5SD above the mean and were excluded from the analyses. The nine participants that reported responding randomly and the four participants who did not provide strategy reports were also excluded from analyses, which left a total of 120 participants. A 3 (Congruency[incongruent, congruent, neutral]) x 3 (Extremity[extreme, moderate, balanced]) x 2 (Time[Time 1, Time 2]) repeated-measures ANOVA was computed on five dependent variables: probability estimates, confidence, response time, personality-sketch AOI fixation time and base-rate AOI fixation time. Results significant at $\alpha = .05$ are reported here. Interactions are decomposed with pairwise comparisons using a Bonferroni correction. Violations of the assumption of sphericity were calculated with Mauchly's test of sphericity and corrected with a Greenhouse-Geisser correction. The correction was only reported when the Greenhouse-Geisser and uncorrected *p*-values were different from each other.

2.6.1. Analysis strategy

The following analyses are divided into four sections. The first section is the analysis of probability estimates to determine that the base-rate neglect phenomenon was replicated. The second section is a test of Hypothesis A that the presumed behavioural indices of conflict detection (confidence ratings, response times, and eye-gaze fixation times) are also sensitive to other cognitive factors, contrasting the predictions of the conflict-detection and coherence-resolution interpretations. The third section is a test of Hypothesis B that reasoners are more explicitly aware of conflict than previously assumed by the conflict-detection interpretation. The fourth and final section is a test of Hypothesis C that base-rate neglect is a function of conflict resolution strategy by determining if conflict resolution strategies map sensibly on to the probability estimates and other dependent measures taken in this experiment.

2.6.2. Base-rate neglect.

The probability estimate data are plotted in Figure 2.2, Figure 2.3, and Figure 2.4. For trials where the category probability to be estimated is the smaller value in the ratio (e.g., "What is the probability that Jack is a lawyer?" when the base-rate ratio has 5 lawyers and 995 engineers), probability estimates were subtracted from 100 to rescale responses so that higher values represent answers closer to the base-rate probability.

The base-rate neglect phenomenon was expected to be found: probability estimates for incongruent problems were predicted to be further from the base-rate probability than estimates for congruent problems. The effect of congruency was significant, $F(2,238)=330.33$, $p<.001$, $\eta_p^2=0.74$. Probability estimates for incongruent trials ($M=37.68$, $SD=13.83$) were further from the base-rate probability than estimates for congruent trial ($M=73.9$, $SD=9.96$; $t(119)=6.63$, $p<.001$), replicating the standard base-rate neglect phenomenon.

Based on previous research (De Neys & Glumicic, 2008; Newman et al., 2017), it was expected that probability estimates for neutral problems would be closer to the base-rate probability than estimates for incongruent problems but also further from the base-rate probability than estimates for congruent problems. As expected, estimates for neutral problems ($M=59.67$, $SD=9.2$) were closer to the base-rate probability than incongruent estimates but further from the base-rate probability than congruent estimates (ts$(119)>18.1$, ps$<.001$).

The main effect of time, $F(1,119)=34.14$, $p<.001$, $\eta_p^2=0.22$, was qualified by a Time x Congruency interaction, $F(2,238)=18.98$, $p<.001$, $\eta_p^2=0.14$. The same pattern of congruency was found at both Time 1 and Time 2: congruent estimates were closer to the base-rate probability than neutral estimates, which were closer to the base-rate probability than incongruent estimates ($ts(119)>9.65$, $ps<.001$). The interaction occurred because, on average, estimates moved closer to the base-rate probability at Time 2 for congruent trials (+7.05; $t(119)=8.46$, $p<.001$) and neutral trials (+4.02; $t(119)=5.16$, $p<.001$), but not for incongruent trials (-1.02; $t(119)=0.85$, $p=.399$). This result is consistent with previous research with the two-response paradigm and the base-rate task that Time 2 probability estimates are closer to the statistically correct probability (Newman et al., 2017; Thompson et al., 2011).

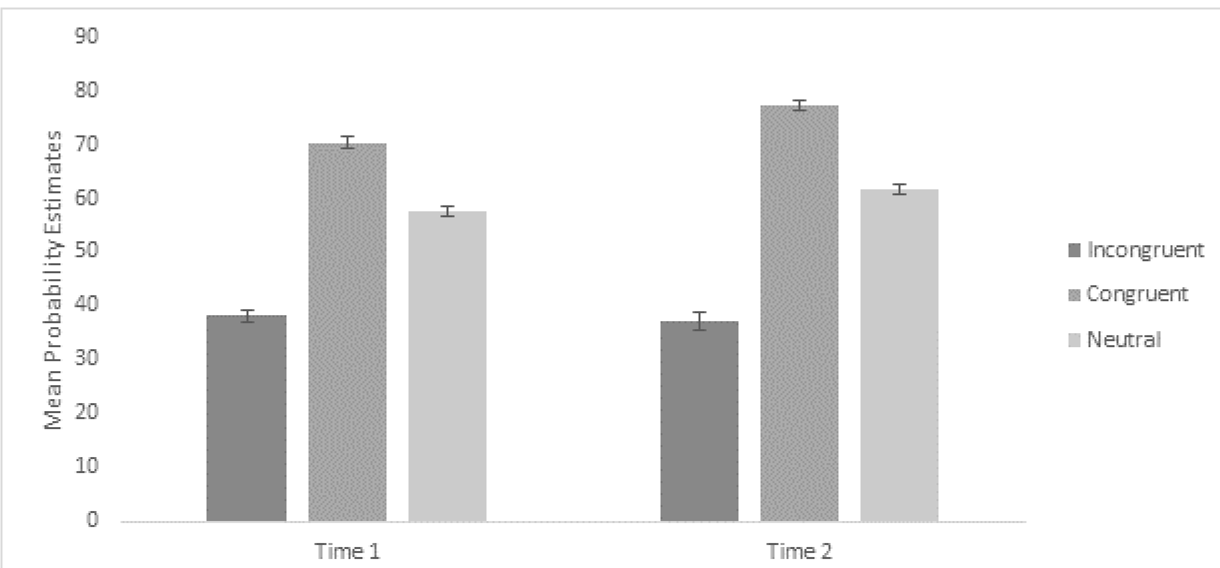

Figure 2.2. Mean probability estimates for Experiment 1 as a function of Time x Congruency. Error bars represent standard errors.

An effect of extremity was also predicted: probability estimates would be highest for extreme ratios and lowest for balanced ratios. The extremity effect was observed, $F(2,238)=91.06$, $p<.001$, $\eta_p^2=0.43$. Probability estimates were highest when ratios were extreme ($M=64.98$, $SD=13.14$) and estimates for moderate ratios ($M=55.87$, $SD=7.95$) were also higher than for balanced ratios ($M=50.35$, $SD=6.46$; ts(119)>8.5, ps<.001), demonstrating that reasoners were sensitive to the differences in base-rate extremity.

The effect of extremity also interacted with time, $F(2,238)=11.83$, $p<.001$, $\eta_p^2=0.09$. The same pattern of the effect of extremity was observable at both Time 1 and Time 2: extreme estimates were higher than moderate estimates, which in turn were higher than balanced estimates (ts(119)>4.82, ps<.001). Overall, probability estimates were closer to the base-rate probability at Time 2 than at Time 1 for extreme base-rates (+6.9; $t(119)=5.54$, $p<.001$) and moderate base-rates (+2.16; $t(119)=2.91$, $p=.004$), but only marginally closer for balanced base-rates (+1.06; $t(119)=1.78$, $p=.077$). In general, this is consistent with previous research that probability estimates are closer to the statistically correct base-rate probability when more time is available to respond (Newman et al., 2017).

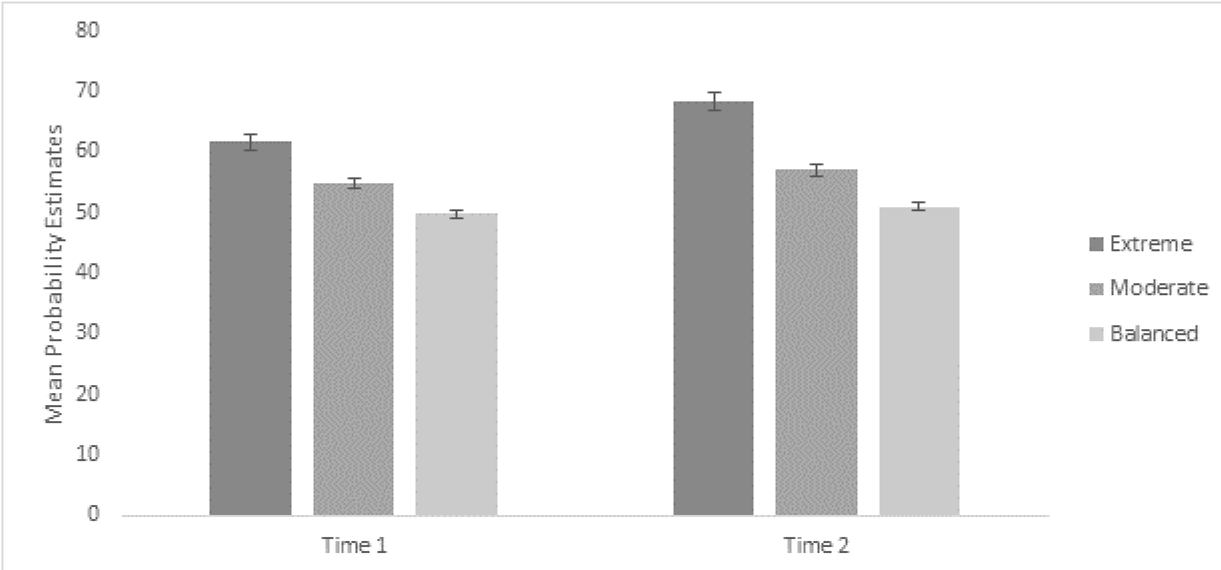

Figure 2.3. Mean probability estimates for Experiment 1 as a function of Time x Extremity. Error bars represent standard errors.

The Congruency x Extremity interaction was significant, $F(4,476)=5.3$, $p<.001$, $\eta_p^2=0.04$. The means for this interaction are found in Table 2.1. The extremity effect was present within each congruency condition (ts(119)>2.88, ps<.015). The difference between extreme and moderate estimates was larger in the neutral condition (+13.26; $t(119)=8.52$, $p<.001$) than the incongruent condition (+6.8; $t(119)=3.64$, $p<.001$) and the congruent condition (+7.29; $t(119)=4.79$, $p<.001$). Thus, the effect of the extremity manipulation was largest in the neutral congruency condition where the personality descriptions are uninformative. This result is generally consistent with the prediction that the extremity effect would be found in the congruent and neutral conditions but only for extreme problems in the incongruent condition (Newman et al., 2017). Instead, probability estimates differed across all congruency and extremity conditions. Lastly, the Time x Congruency x Extremity interaction was not significant, $F(4,476)=1.72$, $p=.143$, $\eta_p^2=0.01$.

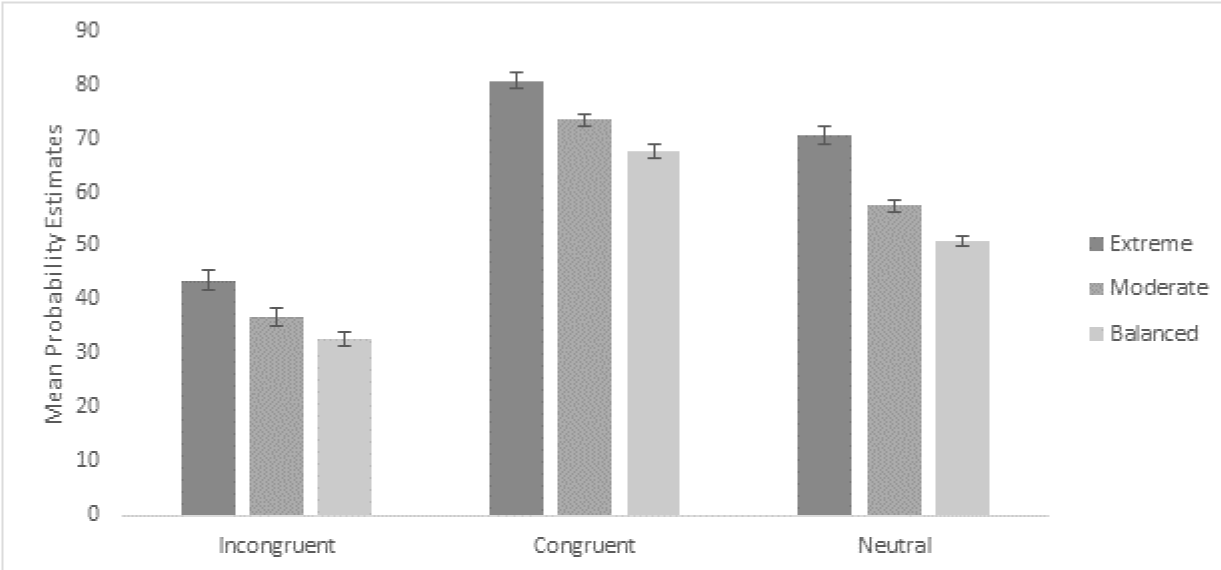

Figure 2.4. Mean probability estimates for Experiment 1 as a function of Congruency x Extremity. Error bars represent standard errors.

Table 2.1. Mean probability estimates by congruency and extremity.

Congruency	Extremity	Mean	SD	N
Incongruent	Extreme	43.61	21.73	120
	Moderate	36.81	16.75	120
	Balanced	32.61	13.38	120
Congruent	Extreme	80.67	15.73	120
	Moderate	73.38	11.71	120
	Balanced	67.53	14.38	120
Neutral	Extreme	70.67	17.97	120
	Moderate	57.41	10.62	120
	Balanced	50.92	11.33	120

2.6.3. Behavioural index sensitivity.

2.6.3.1. Confidence.

The confidence rating data are plotted in Figure 2.5. Based on considerable research on conflict effects in reasoning (De Neys, 2014), confidence for incongruent problems was expected to be lower than for congruent problems. This is one of the hallmark indicators of conflict detection. As predicted, a congruency effect was observed on confidence, $F(2,238)=81.22$,

$p<.001$, $\eta_p^2=0.41$. Incongruent trial confidence (*M*=6.66, *SD*=1.1) was lower than congruent trial confidence (*M*=7.0, *SD*=1.02; $t(119)=6.5$, $p<.001$), replicating the typical effect of conflict on confidence.

Confidence for neutral problems was a critical test of the indices of conflict detection. Lower confidence for incongruent problems (where there is conflict) than the neutral problems is consistent with the conflict-detection interpretation. Conversely, lower confidence for neutral problems than incongruent problems is consistent with the position that there are other cognitive processes during reasoning besides conflict that reduce confidence.

Consistent with past research (Thompson et al., 2011), confidence for neutral trials (*M*=6.2, *SD*=1.07) was lower than both incongruent and congruent trials (*t*s(119)>6.5, *p*s<.001). This result is inconsistent with the conflict-detection interpretation and suggests that lower confidence for incongruent problems (relative to congruent problems) may not index conflict detection *per se*, because reduced confidence is also associated with neutral problems. Potentially, for neutral problems, lack of coherence or relative uncertainty may also function to reduce confidence. If this is the case, one cannot be sure that lower confidence for incongruent problems relative to congruent problems is not also explained (to some degree) by these other cognitive phenomena.

The extremity effect was expected to interact with congruency; varying the base-rate ratio extremity varies the degree of conflict in the problems. Under the conflict detection interpretation, the effect of extremity on confidence should be observable for incongruent problems (i.e., the conflict problems), strongest when the base-rate ratio is extreme, and not present in the congruent and neutral conditions. Alternatively, if confidence is also an index of coherence or uncertainty, extremity would also influence confidence on congruent problems.

The coherence of the base-rate ratio and personality descriptions would serve to increase confidence.

The Congruency x Extremity interaction was significant, $F(4,476)=5.89$, $p<.001$, $\eta_p^2=0.05$. Means for the interaction are presented in Table 2.2. Crucially, contrary to the conflict-detection interpretation, confidence did not differ between extremity conditions for incongruent problems, $F(2,238)=1.19$, $p=.306$, $\eta_p^2=0.01$. As a measure of conflict detection, confidence was not sensitive to the degree of conflict between base-rate ratio and personality description for incongruent problems. In the neutral condition, confidence did not differ between extremity conditions either, $F(2,238)=2.8$, $p=.063$, $\eta_p^2=0.02$.

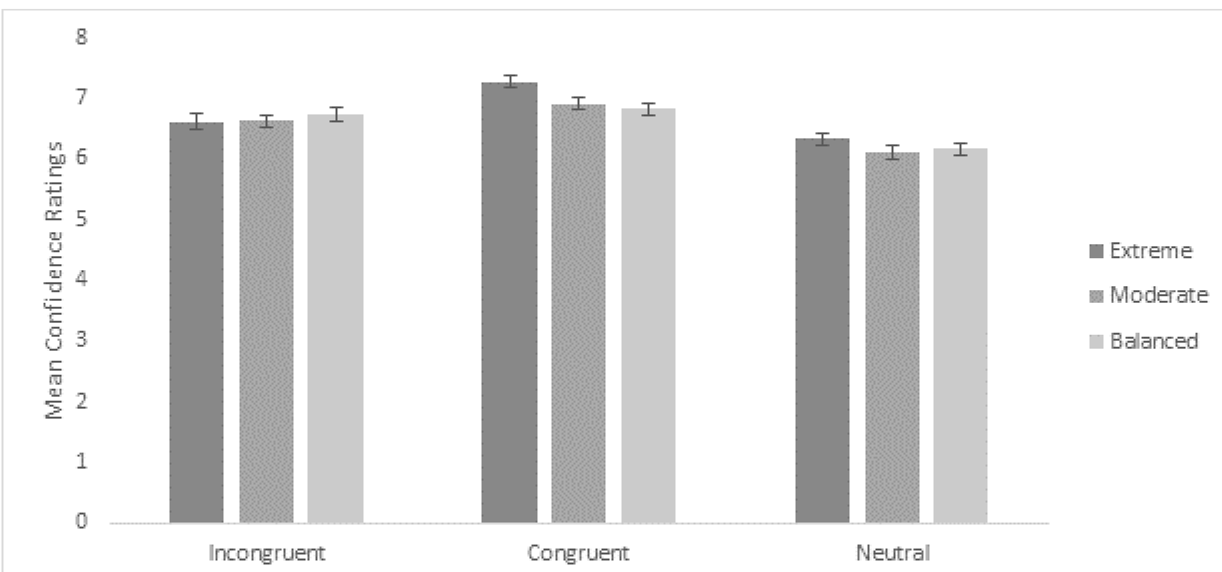

Figure 2.5. Mean confidence ratings for Experiment 1 as a function of Congruency x Extremity. Error bars represent standard errors.

Table 2.2. Mean confidence ratings by congruency and extremity.

Congruency	Extremity	Mean	SD	N
Incongruent	Extreme	6.618	1.283	120
	Moderate	6.621	1.226	120
	Balanced	6.739	1.195	120
Congruent	Extreme	7.275	1.069	120
	Moderate	6.907	1.145	120
	Balanced	6.819	1.215	120
Neutral	Extreme	6.324	1.210	120
	Moderate	6.104	1.272	120
	Balanced	6.158	1.221	120

The source of the interaction was an effect of extremity on confidence in the congruent condition, $F(2,238)=17.71$, $p<.001$, $\eta_p^2=0.13$. Extreme trials had higher confidence ratings than moderate trials ($t(119)=4.37$, $p<.001$) and balanced trials ($t(119)=5.39$, $p<.001$). The moderate trials did not show greater confidence ratings than balanced trials ($t(119)=1.17$, $p=.732$).

Post-hoc *t*-tests indicated that confidence was higher for congruent problems than incongruent problems when base-rate ratios were extreme (+0.66; $t(119)=6.51$, $p<.001$) and moderate (+0.29; $t(119)=3.33$, $p=.001$), but not balanced (+0.08; $t(119)=0.95$, $p=.344$). The predominant conflict-detection interpretation (De Neys, 2014) of this pattern of confidence is that reasoners were sensitive to the degree of conflict in the problems: detecting the conflict reduced their confidence. However, the Congruency x Extremity interaction indicates that the effect of extremity was not present for incongruent problems (nor neutral problems). Instead, it was present for congruent, non-conflict problems. The "conflict" effect was only present for the *non-conflict* congruent problems, strongly suggesting that the typical effect of conflict on confidence is a function of the coherence of information (Koriat, 2012). The conflict-detection interpretation may have been a fundamental misattribution of the effect of conflict, where the assumption was that the effect negatively influenced confidence on conflict trials relative to non-conflict. Instead, coherence may have a positive effect on non-conflict trials relative to conflict trials.

Lastly, the time effect was also significant, $F(1,119)=283.41$, $p<.001$, $\eta_p^2=0.7$, but did not interact with the other factors. Overall, confidence at Time 2 ($M=7.09$, $SD=0.95$) was higher than confidence at Time 1 ($M=6.15$, $SD=1.11$). An increase in confidence over time is consistent with previous evidence that confidence ratings in the two-response paradigm are generally higher at Time 2 (Thompson et al., 2011).

2.6.3.2. Response Times.

Response time data (in seconds) are plotted in Figure 2.6. Under the conflict-detection interpretation, response time and confidence both index the detection of conflict. Thus, where confidence was expected to be lower (i.e., conflict problems), response times were expected to be longer. Alternatively, the assumption that confidence and response time index the same underlying cognitive processes may not be warranted. If the data pattern of response times is not analogous to the data pattern of confidence, it suggests that these measures index different cognitive factors. As proposed above, lack of coherence reducing fluency or relative feelings of uncertainty may account for increased response times.

The typical effect of conflict on response times was replicated with a congruency effect, $F(2,238)=21.59$, $p<.001$, $\eta_p^2=0.15$, such that responses were faster in the congruent condition ($M=12.11$, $SD=3.54$) than the incongruent condition ($M=12.86$, $SD=4.42$; $t(119)=3.71$, $p<.001$). Thus, on average, responses were slower on conflict trials, one of the hallmark indicators of conflict detection.

Critically, responses were slowest overall: slower in the neutral condition ($M=13.54$, $SD=4.19$) than the congruent condition ($t(119)=7.43$, $p<.001$) and incongruent condition

($t(119)=2.68$, $p=.025$). Therefore, the largest effect of "conflict" (i.e., increased response times) occurred in the non-conflict neutral condition, where response times were longest. This result is also inconsistent with the conflict-detection interpretation that predicts the longest response times would be for incongruent problems. De Neys and Glumicic (2008) found that response times were faster for congruent problems than neutral problems but that even longer response times were observed for incongruent problems. Franssens and De Neys (2009) found no differences in response time between incongruent and neutral problems, both of which were longer than for congruent problems. Our data contradicts these previous findings and suggests that response time is an impure index of conflict detection.

The Congruency x Extremity interaction was not significant for response times, $F(4,476)=1.76$, $p=.134$, $\eta_p^2=0.02$; the means for this interaction are in Table 2.3. If response time and confidence indexed the same cognitive factors, this interaction would be expected to be analogous to the Congruency x Extremity interaction on confidence ratings reported above. This study is the first investigation into the interaction of congruency and extremity on the hallmark indices of conflict detection. The absence of a Congruency x Extremity interaction on response times indicates that confidence and response time may index different underlying cognitive factors. To date, this discrepancy has not been observed in the conflict detection literature. There was power of 0.88 to find a relatively small effect size of $\eta_p^2=0.03$ for the response time Congruency x Extremity interaction. Thus, we had good power to detect this interaction but it was not observed.

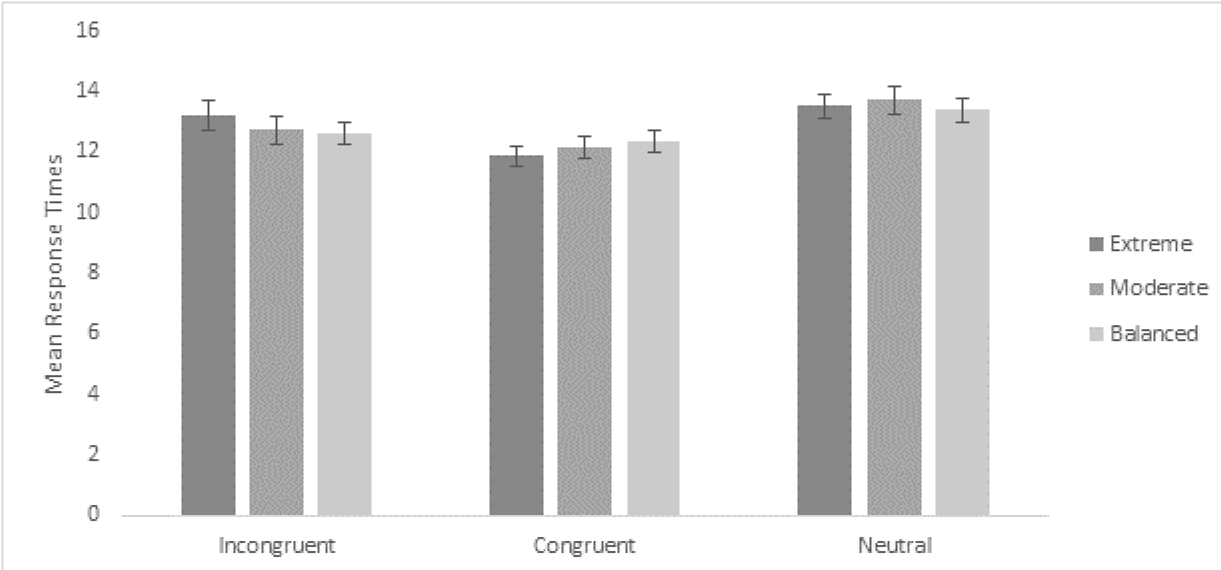

Figure 2.6. Mean response times for Experiment 1 as a function of Congruency x Extremity. Error bars represent standard errors.

Table 2.3. Mean response times by congruency and extremity.

Congruency	Extremity	Mean (ms)	SD	N
Incongruent	Extreme	13213	5417	120
	Moderate	12732	4998	120
	Balanced	12632	4098	120
Congruent	Extreme	11865	3683	120
	Moderate	12131	4101	120
	Balanced	12340	4157	120
Neutral	Extreme	13513	4577	120
	Moderate	13702	4939	120
	Balanced	13389	4329	120

The effect of time was significant, $F(1,119)=6.46$, $p=.012$, $\eta_p^2=0.05$; response times were longer at Time 2 (+1.55; $t(119)=6.55$, $p<.001$). This is unsurprising: participants were instructed to respond with the first thought that came to mind at Time 1 but think carefully at Time 2 and we applied time pressure to Time 1 responses. The time effect also interacted with congruency, $F(1.67,198.7)=4.63$, $p=.016$, $\eta_p^2=0.04$, but patterns were similar between Time 1 and Time 2. At Time 1, incongruent responses ($M=12.0$, $SD=2.63$) were slower than congruent responses

(M=11.66, SD=2.77; $t(119)=3.04$, p=.009); also, neutral responses (M=12.52, SD=2.91) were slower than congruent and incongruent ones (ts(119)>4.03, ps<.001). At Time 2, congruent responses (M=12.56, SD=5.97) were faster than incongruent responses (M=13.72, SD=7.96; $t(119)=3.03$, p=.009) and neutral responses (M=14.55, SD=7.26; $t(119)=6.24$, p<.001), but incongruent and neutral response times did not differ ($t(119)=1.82$, p=.215). The extremity effect was not significant, F<1, nor was the Time x Extremity interaction, $F(2,238)=1.16$, p=.314, $\eta_p^2=0.01$, or Time x Congruency x Extremity interaction, $F(4,476)=1.99$, p=.095, $\eta_p^2=0.02$.

2.6.3.3. Fixation Times.

The purpose of the eye-tracking measures was to measure visual attention. Only the total fixation time (in seconds) in each area of interest per trial are reported[2]. Areas of interest (AOI) were drawn around the base-rate ratio and personality descriptions. The base-rate AOI occupied 9% of the screen and the personality description AOI occupied 27% of the screen.

The total fixation time per trial within each AOI was calculated for analyses. A x 2 (Time[Time 1, Time 2]) x 3 (Congruency[incongruent, congruent, neutral]) x 3 (Extremity[extreme, moderate, balanced]) repeated-measures ANOVA was computed on fixation time for the personality description AOI and base-rate AOI separately. Calibration failed for

[2] Further analyses were conducted on total fixation frequency per AOI and the number of fixation swaps from one AOI to the other. These measures were highly correlated with the fixation time measures (rs>.276; ps<.004) and are not reported here. Analyses of AOI revisits were not significant. Analysis of pupil diameter found a difference between AOIs, $F(1,102)=31.84$, p<.001, $\eta_p^2=0.24$, with larger diameter for the personality sketch AOI. Changes in pupil diameter are related to attention, emotions, arousal, decisions, cognitive load, and working-memory load (see Naber, Frässle, Rutishauser, & Einhäuser, 2013). The multiple differences between our AOIs (size, amount of text and reading required, cognitive and working-memory load) prevent any interpretation of the main effect of AOI. For the base-rate AOI, an effect of time was recorded, $F(1,79)=5.79$, p=.018, $\eta_p^2=0.07$, that interacted with congruency, $F(2,158)=3.53$, p=.032, $\eta_p^2=0.04$, but the effect of congruency was not significant, $F(2,158)=1.16$, p=.316, $\eta_p^2=0.01$. Congruency was not significant at Time 1 (F<1) or Time 2, $F(2,210)=1.12$, p=.33, $\eta_p^2=0.01$. For the personality sketch AOI, all effects were non-significant (ps>.277). The small effect of congruency on base-rate AOI pupil diameter suggests further investigation of pupil diameter changes during attention to probabilistic information or to explore relative weighting of information value (Ariel & Castel, 2014).

four participants; calibration succeeded but the eye-tracker failed to capture their eye-gaze data for three participants.

There is evidence that reasoners increase their attention on the conflicting information on conflict problems (Ball et al., 2006) and to the base-rate ratios in the base-rate task (De Neys & Glumicic, 2008). Therefore, the conflict-detection interpretation predicts that fixation time would be higher for both the personality descriptions and base-rate ratios for incongruent problems than for neutral and congruent problems. The coherence-resolution interpretation is that reasoners would increase their attention to neutral personality descriptions (perhaps as a function of uncertainty). Fixation time that is higher for neutral personality descriptions than for incongruent descriptions provides a challenge to the conflict-detection interpretation.

The personality sketch AOI fixation time data is presented in Figure 2.7. The effect of congruency on fixation time was present on the personality sketch AOI, $F(2,202)=13.24$, $p<.001$, $\eta_p^2=0.12$. Inconsistent with the conflict-detection interpretation, fixation time for congruent ($M=2.76$, $SD=1.49$) and incongruent ($M=2.79$, $SD=1.56$) personality descriptions did not differ ($t(111)=0.43$, $p>.999$). Fixation time was longer on neutral personality descriptions ($M=3.09$, $SD=1.61$) than fixation time on congruent descriptions ($t(111)=4.12$, $p>.001$) and incongruent descriptions ($t(110)=3.76$, $p<.001$). Therefore, reasoners did not respond to the conflict by increasing their attention to the personality descriptions. Instead, relative to congruent and incongruent problems, more attention was focused on the uninformative neutral personality descriptions. Potentially, reasoners were searching for hidden information within the descriptions or attempting to decipher how the neutral descriptions would be relevant to the problems.

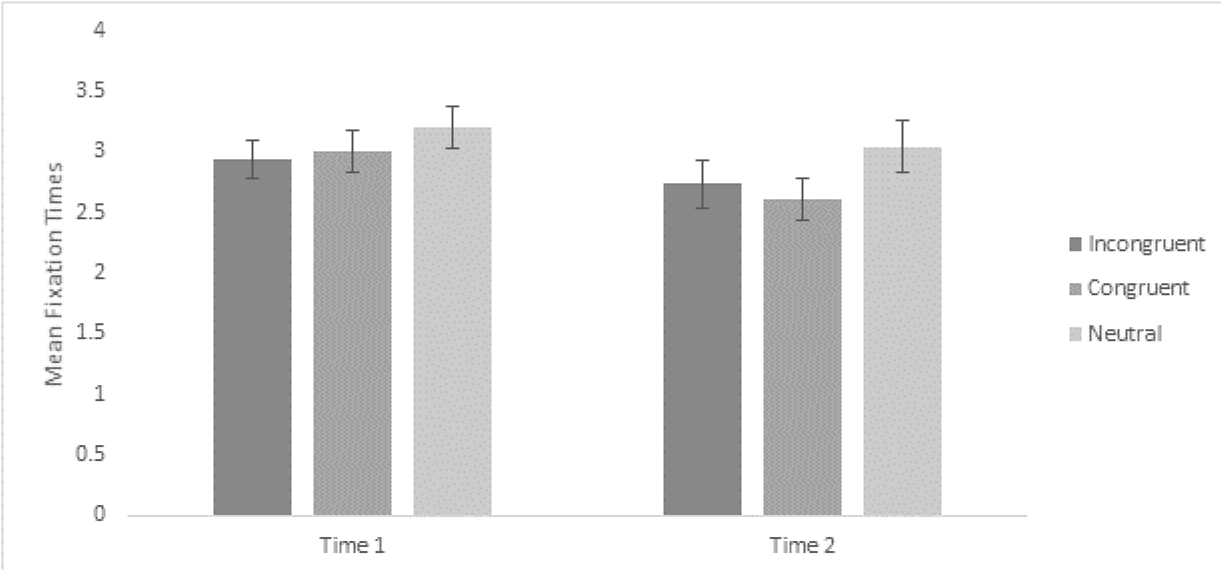

Figure 2.7. Mean fixation times on the personality sketch AOI for Experiment 1 as a function of Time x Congruency. Error bars represent standard errors.

The base-rate AOI fixation time data is presented in Figure 2.8. The congruency effect on fixation time was also present for the base-rate AOI, $F(2,202)=19.18$, $p<.001$, $\eta_p^2=0.16$, qualified by a Time x Congruency interaction, $F(2,202)=7.64$, $p<.001$, $\eta_p^2=0.07$. At Time 1, fixation time on the base-rate ratios for incongruent problems ($M=1.92$, $SD=1.01$) was not different from congruent problems ($M=1.87$, $SD=1.08$; $t(108)=0.89$, $p>.999$) or neutral problems ($M=2.02$, $SD=1.06$; $t(108)=2.18$, $p=.095$) but fixation time on the neutral problem base-rates was higher than for congruent problem base-rates ($t(110)=3.24$, $p=.005$). Thus, inconsistent with the conflict-detection interpretation, the initial reaction to the conflict is not to review the base-rates more for the conflict (i.e., incongruent) problems compared to congruent problems, but rather, more time is spent attending to the neutral base-rates than the congruent base-rates. At Time 2, fixation time on the base-rate ratios was lower for congruent problems ($M=1.53$, $SD=1.24$) than for incongruent problems ($M=1.8$, $SD=1.52$; $t(110)=3.43$, $p=.003$) and neutral problems ($M=1.98$, $SD=1.54$; $t(110)=5.3$, $p<.001$). Fixation time on the base-rate ratios for incongruent and neutral problems did not differ (*t*(110)=2.03, *p*=.135). Thus, at Time 2, more fixation time is spent on the base-rates for incongruent and neutral problems but this result may be deceptive. Post-hoc *t*-tests revealed that fixation times on the base-rates were similar between Time 1 and Time 2 for incongruent problems (-0.07; t(108)=0.52, p=.605) and neutral problems (+0.004; t(110)=0.52, p=.974) but significantly lower at Time 2 for congruent problems (-0.3; t(110)=2.64, p=.01). Therefore, the interaction occurred because reasoners spent less time focusing on the congruent base-rates at Time 2. Contrary to the conflict-detection interpretation, longer fixation times on the incongruent base-rates (relative to congruent base-rates) did not appear to be re-evaluation of the conflicting statistical information in response to conflict but instead, a reduction of attention to congruent base-rates over time.

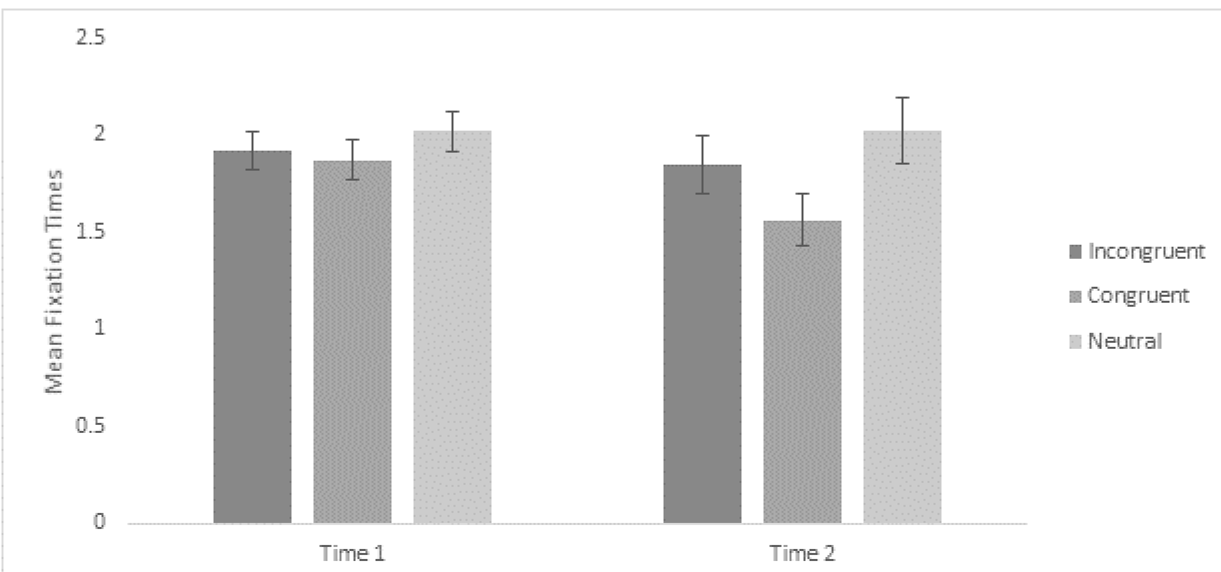

Figure 2.8. Mean fixation times on the base-rate AOI for Experiment 1 as a function of Time x Congruency. Error bars represent standard errors.

2.6.4. Explicit Awareness: Self-Reports.

2.6.4.1. Explicit Awareness Ratings.

Participants responded to three questions intended to index their explicit awareness of conflict in the base-rate problems they had just completed. In general, reasoners responded with relatively high awareness ratings to these questions: Q1 (M=3.61, SD=1.18), Q2 (M=4.45, SD=1.31), and Q3 (M=4.41, SD=1.28). The measures correlated with one another; therefore, a composite score of explicit awareness was computed (M=4.14, SD=0.98) from Q1, Q2, and Q3[3].

Explicit awareness was predicted to correlate with the indices of conflict detection: confidence, response times, and fixation times (see Table 2.4). No correlations were found between explicit awareness and any of the other measures[4] (ρs<.13, ps>.181). Response time and confidence did correlate (r=-.35, p<.001). When this correlation is viewed in isolation, response time and confidence appear to index conflict detection in a manner consistent with the conflict-detection interpretation (De Neys, 2014), but this interpretation overlooks the above evidence that these measures are affected differently by manipulations of congruency and extremity. Despite this, no evidence was found that explicit awareness was related to implicit detection of conflict.

---

[3] Q1 correlated with Q2 (ρ=.39, p<.001) and Q3 (ρ=.28, p=.018), but Q2 and Q3 did not correlate significantly (ρ=.11, p=.37). The subsequent analyses were run on each question separately, in addition to the composite score. These analyses did not differ from the composite score analyses and are not reported here.

[4] Correlations were also computed between the composite score and several of the confidence, response time, and fixation time conditions: Time 1, Time 2, averaged across Time, incongruent, congruent, neutral, a difference score of congruent minus incongruent, and averaged across all conditions. None of these comparisons were significant. Only the variables average across all conditions are reported in Table 3.

Table 2.4. Spearman correlations between the implicit and explicit measures of conflict.

		Explicit Awareness	Confidence	Response Time	Base Rate Fixation Time	Description Fixation Time
Explicit Awareness	Spearman's ρ	—	0.021	-0.015	0.084	0.126
	p-value	—	0.819	0.874	0.375	0.181
Confidence	Spearman's ρ		—	-0.330 ***	-0.079	0.013
	p-value		—	< .001	0.402	0.889
Response Time	Spearman's ρ			—	0.255 **	0.243 **
	p-value			—	0.006	0.009
Base Rate Fixation Time	Spearman's ρ				—	0.639 ***
	p-value				—	< .001
Description Fixation Time	Spearman's ρ					—
	p-value					—

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

2.6.4.2. Resolution Strategy.

One novel feature of this study was to determine whether reasoners could explicitly report the general strategy they adopted to solve the base-rate problems. At the end of the task, participants reported their strategy. Participants that reported relying mostly on the ratios (numbers) or mostly on the descriptions (personality sketches) were placed into the Numbers and Sketches groups, respectively. Those who reported taking a case-by-case approach were sorted into the Cases group. Any participants who reported multiple strategies were placed in the Cases group; participants who reported choosing their answers randomly were excluded from the analyses. Participants that chose to enter a description of their strategy were categorized accordingly. For example, the following description was categorized as a Cases strategy:

> "My decision would vary depending on the [sic] personal sketches or numbers for each case. I would rely [on] one more than the other."

In total, the reported strategies are as follows: Numbers strategy (*N*=26, 21.7%), Sketches strategy (*N*=53, 44.2%), Cases strategy (*N*=41, 34.2%). The proportion of participants reporting a Numbers strategy is consistent with the interpretation that cognitive “biases” observed at the aggregate level obscure a sizable minority of reasoners who prefer rule-based reasoning (Newman et al., 2017) or individual differences in conflict resolution between belief-based and rule-based information (Pennycook & Thompson, 2012). To ascertain the ability of reasoners to (at least retrospectively) explicitly identify their overall approach to solving the reasoning problems, a 3 (Congruency[incongruent, congruent, neutral]) x 3 (Extremity[extreme, moderate, balanced]) x 2 (Time[Time 1, Time 2]) x 3 (Strategy[numbers, sketches, cases]) mixed-design ANOVA was computed for each of the five previous dependent measures (probability estimates, confidence, response times, and fixation times on the personality description and base-rate ratio separately).

2.6.4.2.1. Probability estimates.

It was expected that those who relied primarily on the base-rate ratios would give probability estimates closer to the base-rate probability than those who relied on the personality descriptions. Presumably, those who took a case-by-case approach would perform like an average of Numbers and Sketches strategies. These data are plotted in Figure 2.9 and Figure 2.10.

An effect of strategy was found, $F(2,117)=25.27$, $p<.001$, $\eta_p^2=0.3$. As predicted, the Numbers group gave probability estimates closer to the base-rate probability (M=63.39, SD=5.47) than the Sketches group (M=53.83, SD=5.01; $t(77)=7.1$, $p<.001$) and Cases group (M=57.25, SD=6.42; $t(65)=4.36$, $p<.001$); Sketches group estimates were also further from the base-rate probability than Cases group estimates ($t(92)=2.93$, $p=.012$). Thus, the reported conflict resolution strategies mapped sensibly onto the probability estimate data: those who reported relying on the ratios gave estimates closest to the base-rate probability, while those who reported relying on the personality descriptions gave estimates furthest from the base-rate probability. The case-by-case strategy gave estimates in between the other two strategies.

A Strategy x Congruency interaction was also found, $F(4,234)=11.93$, $p<.001$, $\eta_p^2=0.17$. The coherence-resolution interpretation predicted that reasoners would be able to accurately report their strategy to resolve the conflict in the base-rate problems. Crucially, in the incongruent condition, probability estimates between the strategy groups differed, $F(2,117)=28.94$, $p<.001$, $\eta_p^2=0.33$. For incongruent problems, the Numbers group gave probability estimates closer to the base-rate probability ($M=50.01$, $SD=11.39$) than Sketches ($M=29.75$, $SD=11.18$; $t(77)=7.42$, $p<.001$) and Cases groups ($M=40.11$, $SD=11.7$; $t(65)=3.46$, $p=.002$). The Cases group also gave probability estimates closer to the base-rate probability than the Sketches group ($t(92)=4.37$, $p<.001$). Thus, the effect of strategy was present for incongruent (i.e., conflict) problems.

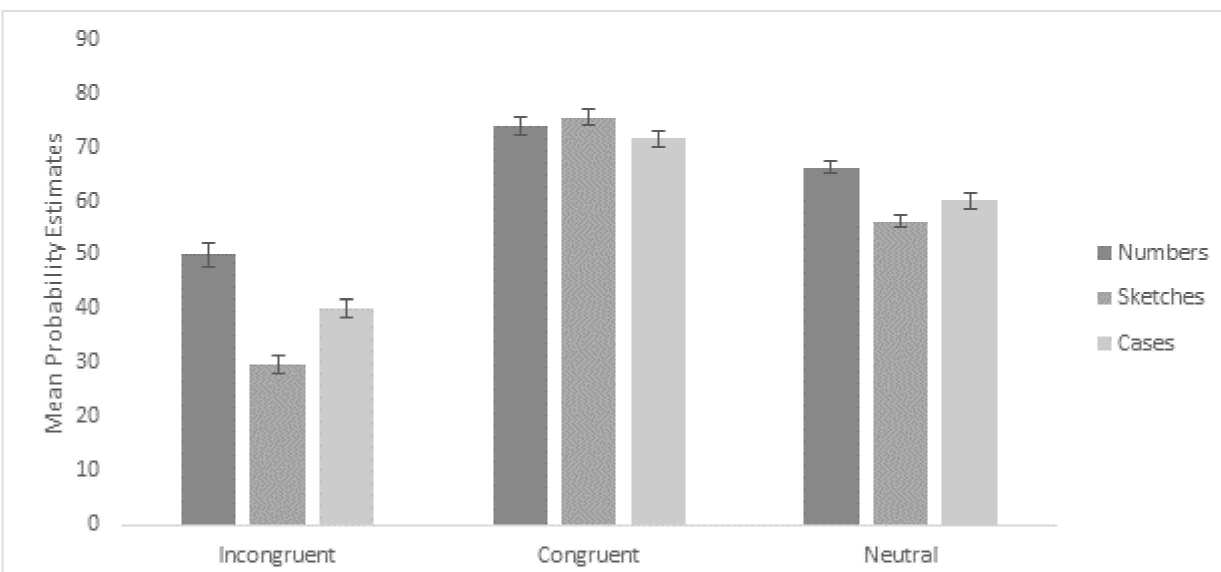

Figure 2.9. Mean probability estimates for Experiment 1 as a function of Strategy x Congruency. Error bars represent standard errors.

For neutral problems, the effect of strategy was also significant, $F(2,117)=12.54$, $p<.001$, $\eta_p^2=0.18$: the Numbers group gave probability estimates closer to the base-rate probability ($M=66.25$, $SD=6.24$) than the Sketches ($M=56.18$, $SD=8.19$; $t(77)=5.0$, $p<.001$) and Cases groups ($M=60.0$, $SD=9.79$; $t(65)=2.96$, $p=.011$), but the Sketches and Cases groups did not differ on their probability estimates ($t(92)=2.18$, $p=.093$). The Numbers group ($M=73.91$, $SD=8.02$), Sketches group ($M=75.55$, $SD=11.09$), and Cases group ($M=71.64$, $SD=9.29$) did not differ on probability estimates for congruent problems, $F(2,117)=1.81$, $p=.169$, $\eta_p^2=0.03$.

In sum, the pattern of probability estimates sensibly mapped on to the self-reported strategies. Probability estimates for congruent and incongruent problems differed the least in the Numbers group (+23.9; $t(25)=7.9$, $p<.001$) and the most in the Sketches group (+45.8; $t(52)=16.76$, $p<.001$). The difference in the Cases group was greater than the Numbers group but less than the Sketches group (+31.53; $t(40)=11.58$, $p<.001$). Therefore, the self-reported strategy groups differed for their probability estimates on conflict (i.e., incongruent) problems relative to non-conflict (i.e., congruent) problems. In other words, these groups differed on their strategy to resolve the conflict and could accurately report on the general strategy they adopted to solve the problems.

Strategy also interacted with extremity, $F(4,234)=13.85$, $p<.001$, $\eta_p^2=0.19$[5]. In the extreme base-rate condition, probability estimates differed between all the groups, $F(2,117)=25.75$, $p<.001$, $\eta_p^2=0.31$. The Numbers group gave probability estimates closer to the base-rate probability ($M=77.22$, $SD=12.08$) than the Sketches ($M=58.32$, $SD=9.4$; $t(77)=7.15$, $p<.001$) and Cases groups ($M=65.83$, $SD=12.27$; $t(65)=4.12$, $p<.001$); the Cases group also gave

[5] Qualified by a Time x Extremity x Strategy interaction, $F(3.58,209.66)=4.28$, $p=.003$, $\eta_p^2=0.07$. The interaction with time occurred because the Numbers group gave probability estimates closer to the base-rate probability than the Sketches and Cases groups for only extreme base-rates at Time 1 ($ts>4.28$, $ps<.001$); the reported Extremity x Strategy interaction pattern emerged at Time 2.

probability estimates closer to the base-rate probability than the Sketches group ($t(92)=3.27$, $p=.004$). For moderate base-rates, the effect of Strategy was also significant, $F(2,117)=15.14$, $p<.001$, $\eta_p^2=0.21$. The Sketches group ($M=52.09$, $SD=7.28$) gave probability estimates further from the base-rate probability than the Numbers; ($M=60.98$, $SD=6.35$; $t(77)=5.2$, $p<.001$) and Cases groups ($M=57.5$, $SD=7.43$; $t(92)=3.64$, $p=.001$). The probability estimates of the Numbers and Cases groups did not differ on moderate problems ($t(65)=1.94$, $p=.164$). For balanced base-rates, the effect of strategy was marginally significant, $F(2,117)=3.09$, $p=.049$, $\eta_p^2=0.05$. The Numbers group ($M=51.97$, $SD=3.727$) gave probability estimates that did not differ from the probability estimates given by the Sketches group ($M=51.06$, $SD=6.93$; $t(77)=0.6$, $p>.999$) or Cases group ($M=48.41$, $SD=6.83$; $t(65)=2.24$, $p=.082$). Probability estimates for the Sketches and Cases group also did not differ ($M=75.55$, $SD=11.09$; $t(92)=2.01$, $p=.141$). The group most sensitive to the extremity manipulation is the group that reported a strategy of reliance upon the ratios, and the group least sensitive is the group that reported primarily relying on the personality sketches. In sum, the probability estimates match closely the strategy reports given by participants, suggesting that reasoners can give accurate retrospective assessments of their conflict resolution strategy.

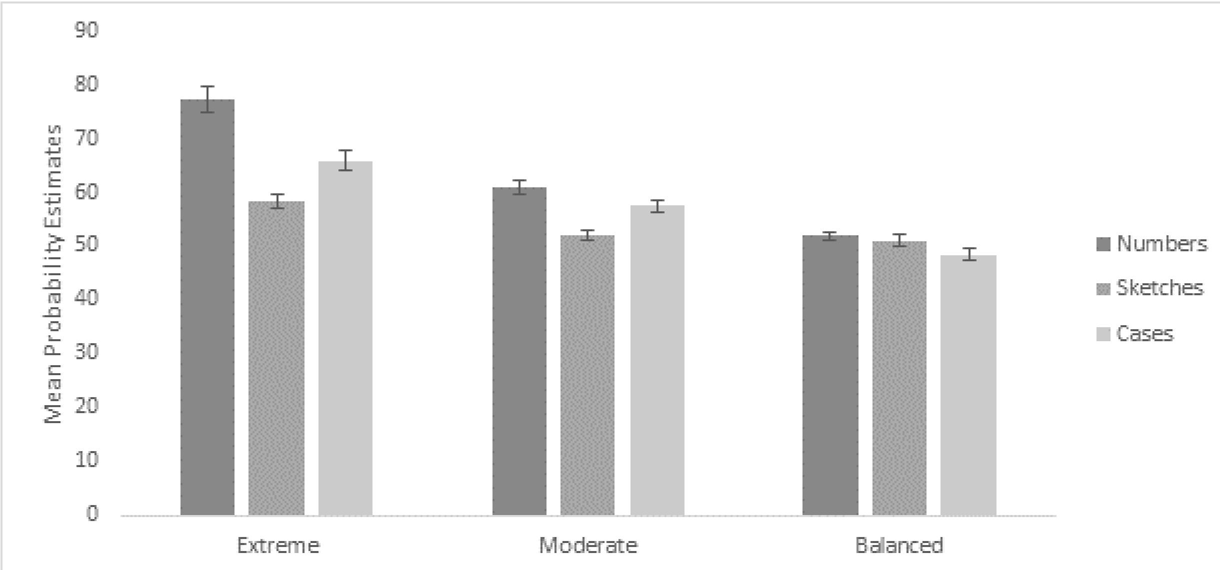

Figure 2.10. Mean probability estimates for Experiment 1 as a function of Strategy x Extremity. Error bars represent standard errors.

2.6.4.2.2. Confidence.

There was no effect of self-reported strategy on confidence ($F<1$), but there was a Congruency x Strategy interaction, $F(4,234)=5.76$, $p<.001$, $\eta_p^2=0.09$. The data are presented in Figure 2.11. The test of the strategy effect on incongruent problems was significant, $F(2,117)=3.07$, $p=.05$, $\eta_p^2=0.05$, but pairwise tests found no differences between each group for confidence. For incongruent problems, the Numbers group reported confidence ($M=6.44$, $SD=1.08$) that did not differ from the reported confidence in the Sketches group ($M=6.93$, $SD=1.08$; $t(77)=1.92$, $p=.172$) or the Cases group ($M=6.44$, $SD=1.08$; $t(65)=0.08$, $p>.999$). The Sketches and Cases groups also reported confidence that did not statistically differ ($t(92)=2.18$, $p=.094$). Post-hoc *t*-tests found that the Sketches group did show higher confidence on incongruent problems than the Cases group ($t(92)=2.18$, $p=.032$) and marginally higher than the Numbers group ($t(77)=1.92$, $p=.058$). These data suggest there is some difference in sensitivity to conflict (as indexed by confidence) between strategies, where those who report relying on a

Sketches strategy are less sensitive to conflict. The Numbers group (*M*=6.83, *SD*=1.14), Sketches group (*M*=7.17, *SD*=1.09) and Cases group (*M*=6.89, *SD*=0.84) did not differ on their confidence for congruent problems, $F(2,117)=1.39$, $p=.253$, $\eta_p^2=0.02$. For neutral problems, it was predicted that the Sketches group would display the lowest confidence, but reported confidence from the Numbers group (*M*=6.25, *SD*=1.25), Sketches group (*M*=6.15, *SD*=1.05), and Cases group (*M*=6.22, *SD*=1.01) did not differ ($F<1$).

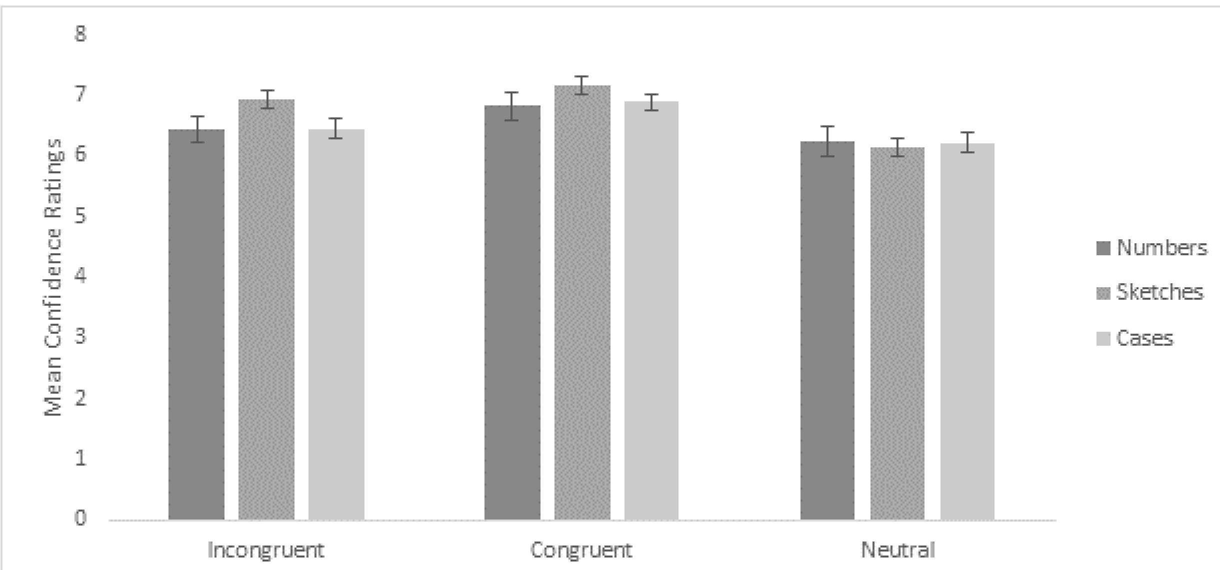

Figure 2.11. Mean confidence ratings for Experiment 1 as a function of Strategy x Congruency. Error bars represent standard errors.

There was a Time x Strategy interaction, $F(2,117)=4.37$, $p=.015$, $\eta_p^2=0.07$. However, confidence did not differ between groups at Time 1, $F(2,117)=1.94$, $p=.149$, $\eta_p^2=0.03$, or at Time 2, ($F<1$). Time 1 confidence in the Numbers (*M*=5.93, *SD*=1.28), Sketches (*M*=6.37, *SD*=1.1), and Cases (*M*=6.01, *SD*=0.97) groups was lower than at Time 2 for each group: Numbers (*M*=7.08, *SD*=1.04; $t(25)=7.47$, $p<.001$), Sketches (*M*=7.14, *SD*=1.0; $t(52)=10.43$, $p<.001$), and Cases (*M*=7.03, *SD*=0.84; $t(25)=12.55$, $p<.001$). The interaction occurred because the difference was similar in the Numbers (+1.15) and Cases (+1.02) groups, but smaller in the Sketches group (+0.77) relative to the other groups.

2.6.4.2.3. Response time.

For response times, the prediction was that the Cases group would take the longest to respond. Instead, the self-reported strategy groups did not differ, $F(2,117)=2.38$, $p=.097$, $\eta_p^2=0.04$, but strategy did interact with time, $F(2,117)=4.35$, $p=.015$, $\eta_p^2=0.07$. At Time 1, response times for the Numbers group ($M=11.79$, $SD=3.14$), Sketches group ($M=12.25$, $SD=2.69$), and Cases group ($M=11.99$, $SD=2.33$) did not differ ($F<1$), presumably because of the time pressure applied to Time 1 responses.

At Time 2, response times differed as a function of strategy, $F(2,117)=3.71$, $p=.027$, $\eta_p^2=0.06$. Overall, the Cases strategy had the longest response times ($M=15.62$, $SD=7.58$), which were longer than the Sketches strategy ($M=11.94$, $SD=4.47$; $t(92)=2.72$, $p=.023$). The Numbers group response times ($M=13.85$, $SD=6.72$) were not different from the Sketches ($t(77)=1.23$, $p=.665$) or Cases strategy response times ($t(65)=1.08$, $p=.848$). The effect of strategy on response times was found in the predicted pattern: taking a case-by-case approach to conflict resolution was more time consuming than relying primarily on one source of information (i.e., numbers or sketches strategies).

There was also a Strategy x Extremity interaction, $F(4,234)=2.82$, $p=.026$, $\eta_p^2=0.05$; these data are presented in Figure 2.12. Response times for the Numbers group ($M=12.53$, $SD=4.28$), Sketches group ($M=12.22$, $SD=3.58$), and Cases group ($M=13.91$, $SD=4.4$) did not differ for the extreme base-rates, $F(2,117)=2.17$, $p=.119$, $\eta_p^2=0.04$. For the balanced base-rates, the Numbers group ($M=12.61$, $SD=3.93$), Sketches group ($M=12.26$, $SD=3.42$), and Cases group ($M=13.59$, $SD=3.94$) also did not differ on response times, $F(2,117)=1.52$, $p=.223$, $\eta_p^2=0.03$. For moderate base-rates, the strategy effect was significant, $F(2,117)=3.4$, $p=.037$, $\eta_p^2=0.06$. Response times for the Numbers group ($M=13.33$, $SD=4.23$) were not different from Sketches group ($M=11.8$, $SD=3.31$; $t(77)=1.58$, $p=.351$) or Cases group response times ($M=13.92$, $SD=4.74$; $t(65)=0.58$, $p>.999$), but the Sketches group did have shorter response times than the Cases group ($t(92)=2.52$, $p=.039$).

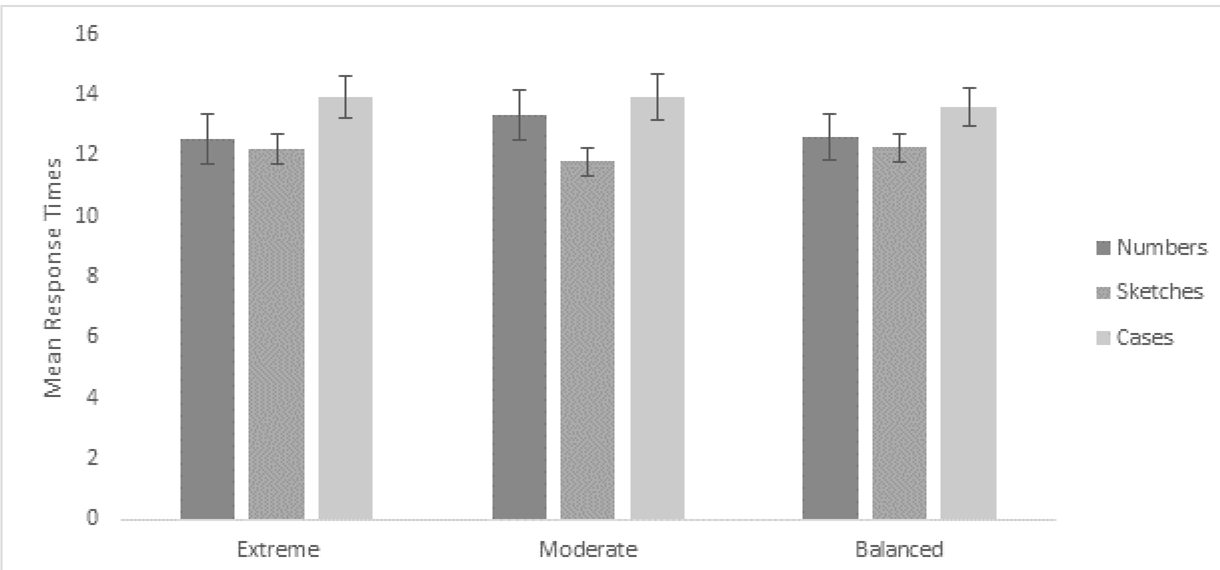

Figure 2.12. Mean response times for Experiment 1 as a function of Strategy x Extremity. Error bars represent standard errors.

2.6.4.2.4. Eye-tracking.

A 2 (AOI[base-rate, personality description]) x 3 (Congruency[incongruent, congruent, neutral]) x 3 (Extremity[extreme, moderate, balanced]) x 2 (Time[Time 1, Time 2]) x 3 (Strategy[numbers, sketches, cases]) mixed-design ANOVA was computed to determine relative differences in fixation time between strategy groups. A main effect of AOI, $F(1,99)=71.73$, $p<.001$, $\eta_p^2=0.42$, was qualified by an AOI x Strategy interaction, $F(2,99)=3.77$, $p=.026$, $\eta_p^2=0.07$. The data are presented in Figure 2.13. For the personality description AOI, the

Numbers group (*M*=2.61, *SD*=1.68), Sketches group (*M*=2.95, *SD*=1.39), and Cases group (*M*=2.77, *SD*=1.58) did not differ on fixation times, *F*<1. Also, fixation time on the base-rate AOI for the Numbers group (*M*=2.05, *SD*=1.07), Sketches group (*M*=1.64, *SD*=0.97), and Cases group (*M*=1.87, *SD*=1.11) did not differ, $F(2,111)=1.4$, $p=.25$, $\eta_p^2=0.03$. Each group fixated less on the base-rate ratios than the personality descriptions: Numbers group ($t(24)=2.33$, $p=.026$), Sketches group ($t(49)=8.36$, $p<.001$), and Cases group ($t(38)=5.18$, $p<.001$). The difference in fixation time between the personality description and base-rate was larger for the Sketches group (+1.31) than the Cases group (+0.9) and Numbers group (+0.56). In other words, this interaction occurred because the ratio of fixation time of personality description to base-rate ratio is smallest in the Numbers group and largest in the Sketches group. This pattern reflects what would be expected of these conflict resolution strategies.

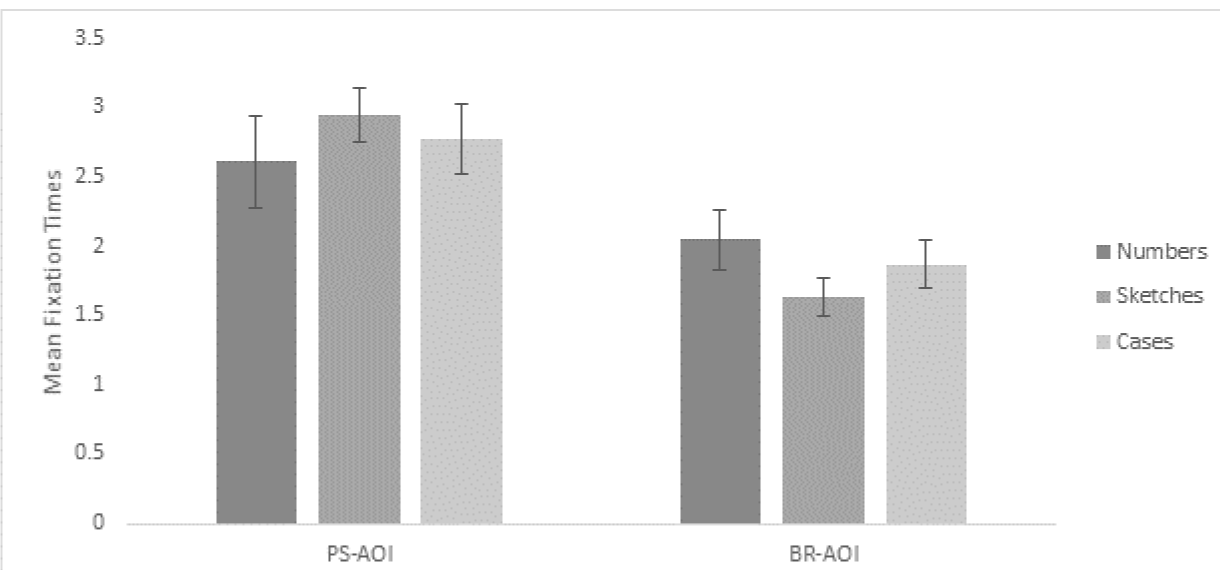

Figure 2.13. Mean fixation times for Experiment 1 as a function of Strategy x AOI. Error bars represent standard errors.

2.6.4.2.5. Explicit awareness.

The Cases group was predicted to report the highest explicit awareness of conflict. A one-way ANOVA was calculated; the effect of strategy was significant, $F(2,117)=3.63$, $p=.029$, $\eta_p^2=0.06$. The Cases group reported explicit awareness ($M=4.46$, $SD=0.9$) that was marginally higher than explicit awareness reports from the Numbers ($M=3.89$ $SD=1.27$; $t(65)=2.38$, $p=.057$) and Sketches groups ($M=4.02$, $SD=0.82$; $t(92)=2.23$, $p=.083$). The Numbers and Sketches groups did not differ on their explicit awareness reports ($t(77)=0.56$, $p>.999$). This result suggests that explicit awareness of conflict may be a function of resolution strategy; the strategy that would weigh the conflicting evidence most equally reported the highest explicit awareness of conflict.

2.7. Discussion

The results of Experiment 1 do not align with the predictions of the conflict-detection interpretation (De Neys, 2012; 2014). Instead, the conflict phenomena are more consistent with a coherence-resolution interpretation. The presumed indices of conflict detection appeared to be sensitive to other cognitive factors than conflict and did not appear to measure the same underlying cognitive processes. Reasoners also reported an explicit awareness of the conflict, which is inconsistent with the conflict-detection interpretation.

The hallmark indicators of conflict detection may have been misinterpreted in the conflict detection literature. For congruent problems relative to incongruent problems, the typical conflict effects of lower confidence, longer response times, and longer fixation times on the conflicting base-rate ratios and personality descriptions replicated the patterns reported in the conflict detection literature (De Neys, 2014). Critically, each of these “conflict” effects (lower confidence, longer response times, longer fixation times) were found to be more pronounced for neutral (i.e., non-conflict) problems, suggesting that these measures are indexing other cognitive factors than purely conflict detection, such as coherence monitoring and conflict resolution. Furthermore, confidence varied as a function of the base-rate ratio extremity for non-conflict (i.e., congruent) problems, not conflict (i.e., incongruent) problems, suggesting that the traditional effect of conflict on confidence is not a relative decrease in the presence of conflict, but rather, a relative increase in the presence of coherent information.

It has been assumed that behavioural measures, such as confidence, response times, and eye-gaze fixation, all index the same underlying cognitive process of conflict detection (De Neys, 2014). Contrary to this interpretation, the interaction of congruency and extremity on confidence ratings was not consistent with the response time and fixation time data. In other words, the presumed conflict detection effects were not consistent across the typical conflict detection indices, suggesting that these measures are indexing different underlying cognitive factors. Potentially, response times and fixation times may index conflict resolution, while confidence indexes detection of the coherence of information. In sum, consistent with Hypothesis A, the measures of confidence, response times, and fixation times appeared to be sensitive to (a) multiple cognitive factors and (b) not necessarily the same cognitive factors as each other.

The base-rate neglect phenomenon was replicated, as was pattern of probability estimates when neutral problems are included in the task (Newman et al., 2017): probability estimates were closest to the base-rate probability for congruent problems and furthest from the base-rate probability for incongruent problems. Estimates for neutral problems were closer to and further from the base-rate probability than incongruent and congruent estimates, respectively. Additionally, the Numbers group provided probability estimates closest to the base-rate probability, while the Sketches group gave estimates furthest from the base-rate probability.

This result is consistent with the interpretation that base-rate neglect is a function of strategy (Pennycook & Thompson, 2012), where the strategy that would lead to the statistically correct response was applied by only a minority of reasoners (in this sample, approximately 20% of participants). In sum, the pattern of probability estimates is consistent with past research and Hypothesis C: the base-rate neglect phenomenon is a function of the resolution strategy adopted by the reasoner.

The retrospective strategy self-reports provided by reasoners were accurate regarding their conflict resolution strategy, mapping sensibly on to all the behavioural measures. As indexed by confidence, the Sketches group appeared least sensitive to conflict, which would be expected given that they value the base-rate ratios the least; the manipulation of conflict is achieved through changing the base-rate ratios. The Cases group took the longest to respond, as they were the group that would have needed the most time to evaluate the evidence and resolve the conflict. The group most explicitly aware of the conflict was the Cases group. The pattern of visual attention also trended towards the expected pattern: less fixation time on base-rate ratios for the Sketches group. Taken together, these data suggest that reasoners are capable of insightful, accurate self-reports regarding their performance and strategy chosen on this reasoning task.

Finally, consistent with Hypothesis B, reasoners reported relatively high explicit awareness of conflict. Surprisingly, this measure did not correlate with any of the typical implicit indices of conflict detection. Two interpretations of these data are available. The first is that the implicit and explicit measures are indexing different cognitive factors (e.g., detection and resolution, respectively). The second is that these measures were not valid assessments of explicit awareness of conflict, despite the apparent success of the strategy self-reports. The goal of Experiment 2 was to address this concern by manipulating explicit awareness, rather than measuring it.

Chapter 3. Experiment 2

The goals of Experiment 2 were twofold. First, the predicted relationship between implicit measures of conflict detection and explicit measures of conflict awareness was not found in Experiment 1. One possibility is that the novel measures of explicit awareness were not sufficiently sensitive to index explicit awareness. Experiment 2 is an attempt to experimentally manipulate the explicit awareness of conflict. Second, the novel measure of conflict resolution strategy from Experiment 1 was found to predict behavioural measures in the predicted manner. Thus, this experiment is also a replication of the strategy effects found on probability estimates, confidence, and response times in Experiment 1.

This study attempted to experimentally increase and decrease explicit awareness with a between-participants manipulation using three groups: an intervention group, a dual-task group, and a control group. A base-rate task was used and problems were presented in two equal blocks of trials. To increase explicit awareness, a retrospective questionnaire (the same set of questions regarding explicit awareness from Experiment 1) was given to the intervention group after they had completed the first block of problems. All three groups completed the same retrospective questionnaire at the end of the task, as well (the intervention group responded to this set of questions twice, in total). In Experiment 1, these questions were given at the end of the task to avoid alerting reasoners to the conflict in the questions. In this case, the intention is to actively notify participants in the intervention group that the problems they are solving contain some degree of conflicting information or competing responses. It was expected that these straightforward queries would alert reasoners to conflict and increase their explicit awareness of conflict in the second half of the task (i.e., Block 2).

To decrease explicit awareness, a dual-task paradigm was used. Concurrent secondary tasks are applied to occupy working-memory resources (Bethell-Fox & Shepard, 1988). This paradigm has been previously applied in the reasoning literature. De Neys (2006) used a dual-task paradigm for conjunction fallacy (Tversky & Kahneman, 1974) and Wason selection (Wason, 1966) reasoning tasks. He found that fewer normatively correct responses were given under working-memory load; incorrect responses also took less time than correct ones. It was concluded that the concurrent task occupied working-memory resources and reduced relative reasoning ability of the participants.

In this study, the goal was to use a secondary task to interfere with the ability to recognize conflict in the problems for the dual-task group. It has been argued that successful conflict detection is a necessary (but not sufficient) prerequisite for correct responding on conflict reasoning tasks (De Neys, 2012; 2014). Therefore, the secondary task may serve to interfere with explicit awareness of conflict, but not necessarily prevent implicit detection of the conflict. A control group was also included as a baseline.

Alongside the block and between-participants manipulations, congruency was also manipulated using congruent, incongruent, and balanced problems. All personality descriptions were informative; no neutral descriptions were used in Experiment 2. Congruent and incongruent problems contained extreme base-rates while balanced problems contained balanced base-rates. Balanced problems were included as a control condition that was expected to be analogous to neutral problems. In neutral problems, the personality description is uninformative; in balanced problems, the base-rate ratio is uninformative.

3.1. Hypotheses

3.1.1. Probability estimates, confidence ratings, and response times.

Probability estimates were expected to be closer to the base-rate probability for congruent problems than incongruent problems; probability estimates for balanced problems were predicted to be higher than incongruent problems but lower than congruent problems (because the base-rate probability is approximately 0.5 in balanced problems, probability estimates are not compared on their proximity to the base-rate probability but instead, their overall magnitude). Similarly, the conflict effects found in Experiment 1 were predicted to be found in Experiment 2. Confidence was expected to be higher for congruent problems than incongruent problems. Also, response times were expected to be longer for incongruent problems than congruent problems. Based on Experiment 1 data for balanced base-rates, balanced problem confidence and response times were expected to be comparable to incongruent problem confidence and response times.

3.1.2. Hypothesis 2A: Explicit awareness of conflict can be increased and decreased.

The between-participants manipulation was expected to influence explicit awareness reports: the intervention group would report higher explicit awareness than the control group. The dual-task group was expected to report lower explicit awareness than the control and intervention groups in both blocks. Also, the intervention group would report higher explicit awareness in the questionnaire at the end of the task after Block 2 than they reported in the questionnaire at the mid-point of the task after Block 1.

3.1.3. Hypothesis 2B: The relationship between base-rate neglect and conflict resolution strategy is robust.

The robust relationship between conflict resolution strategy and probability estimates was also expected to be replicated: for the critical incongruent problems, the Numbers group would give responses closer to the base-rate probability than the Cases group, who in turn would give responses closer to the base-rate probability than the Sketches group. Furthermore, for incongruent problems, the Cases group would have the longest response times and the Sketches group would report the highest confidence. Finally, the Cases group would report the highest explicit awareness of conflict.

3.2. Method

3.2.1. Participants.

One hundred sixty-five participants (104 female, 61 male, M_{age}=22.1 years) from the University of Saskatchewan took part in the study either for partial course credit or from advertisements posted on the campus website (who were compensated CAN $7.50).

3.2.2. Apparatus and Stimuli.

The task was performed on a Microsoft Windows desktop computer using the E-Prime Psychology Software Tools program (Psychology Software Tools, Pittsburgh, PA). Twenty-four base-rate problems were used. Incongruent and congruent trials were consistent with Experiment 1. Instead of neutral problems with uninformative descriptions, balanced base-rates with informative personality descriptions (the same as incongruent and congruent) were used as a control condition. Eight problems from each congruency condition were solved (four in each block). All base-rates in the congruent and incongruent problems were extreme (e.g., 995:5) with four variations (e.g., 996:4, 995:5, 996:4, 997:3). Base-rates in the balanced problems were all approximately 50:50 (e.g., 510:490) with four variations. Problems were presented in two blocks. Each block contained a list of twelve problems; the order of the lists was counterbalanced across participants. Congruency was also counterbalanced across participants. The 24 base-rate problems were tested equally often in each congruency condition: congruent, incongruent, and balanced. Problem content was not repeated throughout the task for any participant. For a full list of base-rate problems, refer to Appendix D.

The same self-report question of conflict resolution strategy from Experiment 1 was given in Experiment 2; this question was presented to all participants at the end of the base-rate task and was not included in the mid-task questionnaire in the intervention group. The same explicit awareness questions from Experiment 1 were also given at the end of the task to all the groups (and also to the intervention group after Block 1). As a further attempt to discriminate those who display low and high explicit awareness of conflict, one of the questions was changed from a 7-point Likert scale rating to a yes-no response option: "Did you notice that sometimes the personality sketch and the numbers suggested different responses?" Participants who responded "no" to this question did not answer the other two explicit awareness questions.

The spatial memory task was a same-different task using a 5x5 grid of squares (adapted from Trbovich & LeFevre, 2003); each trial type was equally likely. In the pre-trial grids, eight of the 25 squares on the grid contained an asterisk and each asterisk could not be adjacent to more than one other asterisk on the grid. This was to avoid the patterns resembling recognizable shapes, such as letters (Trbovich & LeFevre, 2003). The post-trial grids contained yes-no response options below the grid. In same-type trials, post-trial grids were identical to the pre-trial grids. In different-type trials, post-trial grids contained one difference from the corresponding pre-trial grid: one asterisk was moved to one adjacent position on the grid from its initial position (above, below, left, or right). The grid position of each asterisk was randomly generated, as were the asterisk and direction of movement in the post-trial grids. The post-trial asterisk movement randomization was subject to the above rule regarding asterisk proximity.

3.2.3. Procedure.

Base-rate problems were presented in two blocks (ordered randomly within block) in a single-response paradigm. A fixation cross preceded each base-rate problem for two seconds. Participants provided probability estimates (on a scale from 0 to 100) subject to the same time-pressure used in Experiment 1 (white text shifted to red text after 11 seconds elapsed). For each probability estimate, participants entered a rating of their confidence in their response on a nine-point Likert scale (1 – Guessing, 9 – Certain I'm Right). Response times were also recorded for each probability estimate once the Enter key was pressed on the keyboard. Upon completion of all 24 base-rate problems, self-reports were taken from each participant, but none of the filler questions from Experiment 1 were used. The same instructions from Experiment 1 were used but omitted information regarding the two-response paradigm used in Experiment 1. Additional instructions were included for participants in the dual-task group (see below), stressing the importance of the visual memory task.

The trial progression for Experiment 2 is displayed in Figure 3.1. A central fixation cross was presented for two seconds, followed by the base-rate problem for the Time 1 probability estimate. After 11 seconds had elapsed and a response had not been entered, the text colour shifted from white to red. The problem in red text remained on screen until a response was provided.

The dual-task group saw a pre-trial grid (displayed for 1.5 seconds) before each base-rate problem, preceded by a notification to "please remember the following grid." After responding and rating their confidence, participants were presented with a post-trial grid and indicated whether the grids were the same or different as the pre-trial grid with a key press (in free time). The grid presented with each base-rate problem was randomly determined for each participant to prevent pairing of particular grids with particular base-rate problems. The secondary-task instructions are as follows:

> Prior to each problem, an image of a 5x5 grid will be displayed. Some of the cells will contain an asterisk (*), and some will be empty. You are to remember which cells have asterisks in them, as another 5x5 grid will be displayed after you respond to the problem. You will be asked to determine if it is the same grid as the initial grid.
>
> It is VERY IMPORTANT that you be as accurate as possible when judging whether the second grid is the same as the first or not. Do your best to remember the grid. You will respond with the arrow keys on the keyboard.

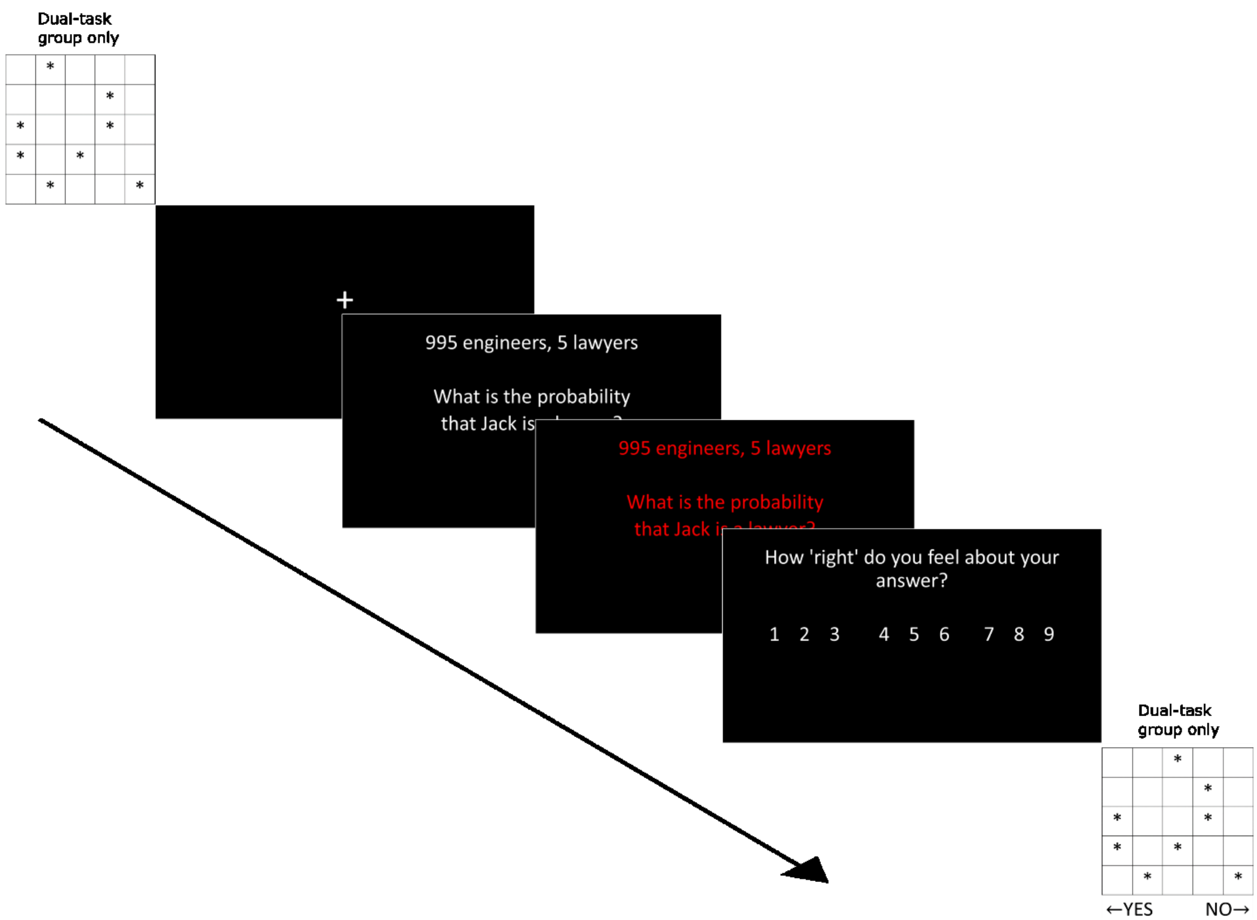

Figure 3.1. Event sequence for Experiment 2 trials (with simplified text for legibility; see Appendix D for problem text). The pre- and post-trial grids were only displayed in the dual-task group.

After completion of the base-rate task, the explicit awareness questions and strategy assessment were taken. The strategy assessment was presented first. The yes-no awareness question was asked next; those who responded “yes” subsequently answered the other two awareness questions. These questions were presented sequentially for all participants. The intervention group had the same set of three explicit awareness questions posed at the midpoint of the task, between Block 1 and Block 2. Completion of the self-report questions concluded the task.

3.3. Results

Mean response time was 12.59 seconds. A mean response time was calculated for each participant; any trial with a response time greater than 2.5SD above their individual mean response time was excluded from the analyses. A total of 2.1% of responses were removed as outliers. The three participants that reported responding randomly and the seven participants who did not provide strategy reports were excluded from analyses, which left a total of 165 participants included in the analyses, divided into three explicit-awareness groups: control (N=52), intervention (N=56), and dual-task (N=57)[6]. Results significant at $\alpha = .05$ are reported here. Interactions are decomposed with pairwise comparisons using a Bonferroni correction. Violations of the assumption of sphericity were calculated with Mauchly’s test of sphericity and corrected with a Greenhouse-Geisser correction; the correction was only reported when the uncorrected and Greenhouse-Geisser p-values differed from each other.

[6] Overall, percentage correct on the spatial working-memory task was 72.8. Also, the strategy groups did not differ on accuracy (F<1): Numbers group (M=72.9, SD=0.09), Sketches group (M=72.2, SD=0.12), and Cases group (M=73.7, SD=0.11).

Two errors were discovered in the computer program for this task. First, a sample of participants (*N*=81) were collected for Experiment 2 but trials were sequentially presented, not randomized. To remedy this, a second sample with proper randomization was collected (*N*=84) to replace the original sample. The two samples of participants were compared as a between-participants factor on all dependent measures. No differences were observed for explicit awareness ratings ($F<1$), confidence (Fs<1.64, ps>.198), and response times (Fs<2.72, ps>.68). An interaction with congruency was found for probability estimates, $F(1.7,268.54)=3.75$, $p=.031$, $\eta_p^2=0.03$. The means are presented in Table 3.1. Probability estimates did not differ for incongruent problems between the first sample and second sample ($t(163)=0.59$, $p=.555$). Similarly, for balanced problems, probability estimates from the first sample did not differ from the second sample ($t(162)=0.12$, $p=.905$). The only difference was found for congruent problems, where probability estimates were higher in the first sample than the second sample ($t(163)=2.4$, $p=.017$). It was reasonable to conclude that this difference was not problematic for the analyses of interest in Experiment 2. Therefore, the samples were combined for the current analyses to increase statistical power. Notably, the same pattern of results reported below is found in analyses using only the second, corrected sample.

Table 3.1. Mean probability estimates by congruency and sample.

	Sample	N	Mean	SD	SE
Incongruent	1	81	41.32	23.761	2.640
	2	84	43.42	21.815	2.380
Congruent	1	81	78.92	13.634	1.515
	2	84	73.81	13.680	1.493
Balanced	1	80	50.80	8.398	0.939
	2	84	50.65	7.550	0.824

The second error was that three of the 24 items used in Experiment 2 contained neutral descriptions instead of the intended informative descriptions. These items were presented equally often in each congruency condition and each block and were removed from the reported analyses. Previously performed analyses that included these items primarily did not differ from the reported analyses below.

3.3.1. Analysis Strategy.

A 3 (Congruency[incongruent, congruent, neutral]) x 2 (Block[Block 1, Block 2]) x 3 (Explicit-Awareness[control, intervention, dual-task]) mixed-design ANOVA was computed for probability estimates, confidence, and response times. The first section is a set of analyses to test for replications of the base-rate neglect phenomenon and the typical effects of conflict: lower confidence and longer response times for incongruent problems relative to congruent problems. The balanced problems served as a control condition; based upon the results of Experiment 1, performance for balanced problems was predicted to be comparable to incongruent problem performance. The second section is a test of Hypothesis 2A: the explicit awareness of conflict would be increased by expressly questioning reasoners about the presence of conflict (i.e., the intervention group) and decreased by occupying working-memory resources with a concurrent task (i.e., the dual-task group). The third and final section is a test for replication of the relationship between self-reported conflict resolution strategy and probability estimates, confidence ratings, response times, and explicit awareness ratings. In other words, the third section is a test of Hypothesis 2B that base-rate neglect and conflict effects in the base-rate task is a function of conflict resolution strategy.

3.3.2. Probability estimates, confidence, and response times.

3.3.2.1. Probability estimates.

For probability estimates, responses were rescaled so higher values corresponded to answers closer to the base-rate probability. The probability estimate data were expected to replicate Experiment 1: estimates for congruent problems would be closer to the base-rate probability than estimates for incongruent problems. Probability estimates for balanced problems were also expected to higher than for incongruent problems but lower than congruent problems. As expected, the effect of congruency on probability estimates mirrored the pattern observed for the analogous conditions in Experiment 1, $F(2,322)=257.01$, $p<.001$, $\eta_p^2=0.62$: estimates were closer to the base-rate probability for congruent trials ($M=76.35$, $SD=13.89$) than incongruent trials ($M=42.07$, $SD=22.43$; $t(164)=19.7$, $p<.001$) and were higher than for balanced trials ($M=50.73$, $SD=7.95$; $t(163)=20.77$, $p<.001$). Balanced estimates were also higher than incongruent estimates ($t(163)=4.82$, $p<.001$).

No main effect of explicit-awareness group was found, $F(2,161)=1.4$, $p=.251$, $\eta_p^2=0.02$, but a Congruency x Explicit-Awareness group interaction was observed, $F(3.43,276.27)=4.0$, $p=.006$, $\eta_p^2=0.05$. The effect of congruency was analyzed within each explicit-awareness group; the data are presented in Table 3.2.

Table 3.2. Mean probability estimates by congruency and explicit-awareness group.

Congruency	Group	Mean	SD	N
Incongruent	Control	36.32	22.865	52
	Intervention	46.78	22.660	55
	Dual-task	42.77	20.993	57
Congruent	Control	79.53	13.244	52
	Intervention	76.66	12.642	55
	Dual-task	73.13	15.100	57
Balanced	Control	50.16	8.484	52
	Intervention	51.49	8.033	55
	Dual-task	50.50	7.436	57

For the control group, the congruency effect was significant, $F(2,102)=114.8$, $p<.001$, $\eta_p^2=0.69$; probability estimates differed between all the congruency conditions (ts(51)>4.18, ps<.001). The effect of congruency was also significant in the intervention group, $F(2,108)=64.61$, $p<.001$, $\eta_p^2=0.55$; probability estimates for congruent problems were closer to the base-rate probability than incongruent problems ($t(55)=10.18$, $p<.001$) and higher than probability estimates for balanced problems ($t(54)=11.37$, $p<.001$). Probability estimates for incongruent and balanced problems did not differ ($t(54)=1.46$, $p=.451$). In the dual-task group, there was also an effect of congruency, $F(2,112)=80.0$, $p<.001$, $\eta_p^2=0.59$. Probability estimates for congruent problems were closer to the base-rate probability than incongruent problems, higher than balanced problems, and balanced estimates were higher than incongruent estimate (ts(56)>2.82, ps<.021).

A one-way ANOVA was calculated for each congruency condition. The effect of explicit-awareness group was significant for incongruent problems, $F(2,162)=3.46$, $p=.034$, $\eta_p^2=0.04$. Overall, the control group gave probability estimates that were lower than the intervention group probability estimates ($t(106)=2.62$, $p=.029$). Probability estimates in the dual-task group did not differ from the control group ($t(107)=1.5$, $p=.406$) or the intervention group ($t(111)=1.16$, $p=.748$). The explicit-awareness groups did not differ on probability estimates for balanced problems, $F<1$, or for congruent problems, $F(2,162)=2.99$, $p=.053$, $\eta_p^2=0.04$.

There was no *a priori* prediction for the effect of the explicit-awareness manipulation on probability estimates. As the Congruency x Explicit-Awareness interaction did not interact with the block factor, $F(3.88,311.99)=2.34$, $p=.057$, $\eta_p^2=0.03$, the effect of explicit-awareness group on probability estimates emerged in Block 1. Potentially, this interaction is a Type I error, explained by chance variance; there is no reason to expect the control and intervention groups to differ significantly in Block 1 where they complete the exact same task.

3.3.2.2. Confidence.

Confidence ratings were expected to replicate the results from Experiment 1, where confidence was rated as higher for congruent problems than incongruent problems. Overall, a congruency effect was observed for confidence ratings, $F(2,324)=7.1$, $p<.001$, $\eta_p^2=0.04$, where confidence for congruent trials ($M=7.11$, $SD=0.99$) was higher than confidence for incongruent trials ($M=6.92$, $SD=1.05$; $t(164)=2.64$, $p=.028$) and balanced trials ($M=6.84$, $SD=1.03$; $t(164)=3.53$, $p=.002$). Confidence did not differ between incongruent and balanced trials ($t(164)=1.09$, $p=.836$). Thus, the congruency effect for incongruent and congruent problems in Experiment 1 was replicated. Notably, confidence on balanced problems was statistically equivalent to incongruent problems; it was the congruent problems where confidence was higher than the other congruency conditions.

3.3.2.3. Response times.

Response times were also expected to replicate the results from Experiment 1, where response times were longer for incongruent problems than for congruent problems. The effect of congruency on response time was not significant, $F(1.91,309.58)=1.19$, $p=.305$, $\eta_p^2=0.01$, but a Block x Congruency interaction was observed, $F(2,324)=3.28$, $p=.039$, $\eta_p^2=0.02$. The congruency effect was not found in Block 1 ($F<1$), but was present in Block 2, $F(1.91,313.48)=4.86$, $p=.009$, $\eta_p^2=0.03$. Overall, Block 2 response times for incongruent problems ($M=12.16$, $SD=3.7$) were longer than Block 2 congruent problem response times ($M=11.58$, $SD=3.17$; $t(164)=3.24$, $p=.004$); response times for balanced problems ($M=12.01$, $SD=4.17$) did not differ from congruent ($t(164)=2.33$, $p=.063$) nor incongruent problem response times ($t(164)=0.68$, $p>.999$). Consistent with the confidence data, balanced problem response times were similar to incongruent problem response times. Also, response times were longer for incongruent problems than congruent problems but this effect only manifested in Block 2. These results suggest that response time is not an effective index of conflict detection, as this effect would be expected to be present at the onset of the task, as well.

An effect of block was found on response times, $F(1,162)=77.11$, $p<.001$, $\eta_p^2=0.32$. Overall, responses were faster in Block 2 ($M=11.92$, $SD=3.4$) than in Block 1 ($M=13.34$, $SD=3.65$); reasoners tended to respond more quickly over time. The block factor also interacted with explicit-awareness group, $F(2,162)=4.04$, $p=.019$, $\eta_p^2=0.05$; the data are presented in Table 3.3. Response times in Block 1 for the intervention, dual-task, and control groups did not differ, $F(2,162)=2.17$, $p=.117$, $\eta_p^2=0.03$. In Block 2, the intervention, dual-task, and control groups also did not differ on response times, $F(2,162)=1.19$, $p=.308$, $\eta_p^2=0.01$. Each group gave faster responses in Block 2 than in Block 1: intervention ($t(55)=5.1$, $p<.001$), dual-task ($t(56)=6.47$, $p<.001$), and control ($t(51)=4.17$, $p<.001$) groups. The interaction occurred because the response time decrease from Block 1 to Block 2 was smaller in the control group (-0.76) than the intervention (-1.81) and dual-task groups (-1.64).

Table 3.3. Mean response times by block and explicit-awareness group.

Block	Group	Mean (ms)	SD	N
Block 1	Control	12813	2248	52
	Intervention	14146	4478	56
	Dual-task	13015	3688	57
Block 2	Control	12054	2058	52
	Intervention	12336	4347	56
	Dual-task	11379	3294	57

3.3.3. Hypothesis 2A: Explicit awareness can be experimentally manipulated.

All participants responded to a yes-no question regarding awareness of conflict. There were 131 "yes" responses (79.4%), which was significantly higher than chance (tested against 50%; $t(165)=25.14$, $p<.001$). Participants in the intervention group also responded to this question between blocks. There were 45 "yes" responses (80.4%) from a total of 56 responses; this value was also higher than chance ($t(55)=15.0$, $p<.001$). Thus, a large majority of participants reported that they were aware that the ratios and descriptions suggested different responses; moreover, they were aware after Block 1 in the intervention group.

If a "yes" response was given, two further questions regarding explicit awareness were presented (this was also the case for the intervention group questionnaire between Block 1 and Block 2). Consistent with Experiment 1, the end-task ratings of explicit awareness were high: Q1 (M=3.84, SD=1.29) and Q2 (M=3.74, SD=1.43). The mid-task ratings in the intervention group were also high: Q1 (M=3.76, SD=1.43) and Q2 (M=3.96, SD=1.35). Thus, a consistent and high proportion of reasoners reported awareness of the conflict.

The primary goal of this experiment was to manipulate explicit awareness of conflict. The prediction was that the intervention group would report higher explicit awareness than the control group and the dual-task group would report lower explicit awareness than the control group. The two end-task measures of explicit awareness were correlated with each other ($\rho=.58$, $p<.001$); the mid-task measures were also correlated with each other ($\rho=.37$, $p=.014$). From these measures, composite mid- and end-task explicit awareness scores were computed.

A one-way ANOVA of explicit awareness was calculated with end-task explicit awareness as the dependent variable. Contrary to the prediction, this effect was not significant, $F(2,128)=0.44$, $p=.646$, $\eta_p^2=0.01$. Explicit awareness ratings did not differ between the intervention (M=3.92, SD=1.21), dual-task (M=3.78, SD=1.32), and control groups (M=3.68, *SD*=1.09), suggesting that neither the intervention nor the dual-task impacted explicit awareness of conflict. Additionally, in the intervention group, the mid-task and end-task scores were not different from each other ($t(37)=0.15$, $p=.880$). This result was surprising, as the intervention was a blatant attempt to notify reasoners of the conflict in the problems. In sum, the attempt to manipulate explicit awareness was unsuccessful (one potential explanation is offered below).

3.3.4. Hypothesis 2B: The effects of resolution strategy will replicate.

The secondary goal of Experiment 2 was to replicate the strategy effects found in Experiment 1. The distribution of strategy reports was similar to Experiment 1; in Experiment 2, 32 participants reported using a Numbers strategy (19.4%), 78 participants reported a Sketches strategy (47.3%), and 55 participants reported a Cases strategy (33.3%). Furthermore, these strategies were similarly distributed across the control, intervention, and dual-task groups, $\chi^2(4,165)=3.14$, $p=.535$. As in Experiment 1, the majority of participants adopted a strategy to rely on the personality descriptions, which explained an overall base-rate neglect phenomenon when probability estimates were analyzed across all participants. To test the effect of strategy, a 3 (Congruency[incongruent, congruent, neutral]) x 2 (Block[Block 1, Block 2]) x 3 (Strategy[Numbers, Sketches, Cases]) mixed-design ANOVA was computed on three dependent variables: probability estimates, confidence, and response times.

3.3.4.1. Probability estimates.

A robust effect of strategy on probability estimates was found in Experiment 1 and was expected to be replicated in Experiment 2. The effect of strategy was significant, $F(2,161)=22.24$, $p<.001$, $\eta_p^2=0.22$. The Numbers group gave probability estimates higher ($M=64.3$, $SD=11.11$) than the Sketches group ($M=51.48$, $SD=7.89$; $t(108)=6.08$, $p<.001$). The Cases group ($M=59.1$, $SD=10.55$) also gave higher than the Sketches group ($t(131)=4.65$, *p*<.001). The Numbers and Cases groups did not differ on their probability estimates ($t(85)=2.1$, $p=.111$).

A Congruency x Strategy interaction was found, $F(4,322)=7.15$, $p<.001$, $\eta_p^2=0.08$. The effect of strategy was interpreted by calculating a one-way ANOVA on probability estimates within each congruency condition (see Figure 3.2). Based on Experiment 1, it was expected that the Numbers group would give responses closest to the base-rate probability and the Sketches group would provide responses furthest from the base-rate probability. For incongruent problems, the effect of strategy was observed, $F(2,162)=19.07$, $p<.001$, $\eta_p^2=0.19$. Consistent with the coherence-resolution interpretation and replicating Experiment 1, the Numbers group ($M=58.23$, $SD=22.96$) gave estimates closer to the base-rate probability than the Sketches group ($M=32.84$, $SD=17.18$; $t(108)=5.87$, $p<.001$) and Cases group ($M=46.72$, $SD=23.42$; $t(85)=2.51$, $p=.039$), while the Cases group gave estimates closer to the base-rate probability than the Sketches group ($t(131)=3.83$, $p<.001$). For congruent problems, the strategy effect was also significant, $F(2,162)=6.69$, $p=.002$, $\eta_p^2=0.08$. The Numbers group gave responses closer to the base-rate probability ($M=81.59$, $SD=14.55$) than the Sketches group ($M=72.43$, $SD=13.37$; $t(108)=3.26$, $p=.004$) and the Cases group ($M=78.78$, $SD=12.74$; $t(85)=2.69$, $p=.023$); the Cases and Sketches groups did not differ on congruent probability estimates ($t(131)=0.94$, $p>.999$). For balanced problems, the strategy effect on probability estimates was not significant, $F(2,161)=2.7$, $p=.071$, $\eta_p^2=0.03$. The Numbers group ($M=52.21$, $SD=6.42$), the Sketches group ($M=49.23$, $SD=8.83$), and the Cases group probability estimates ($M=52.01$, $SD=7.12$) did not differ for balanced problems. In sum, this pattern of data replicates the strategy effect on probability estimates from Experiment 1, but the effect also emerged for congruent problems.

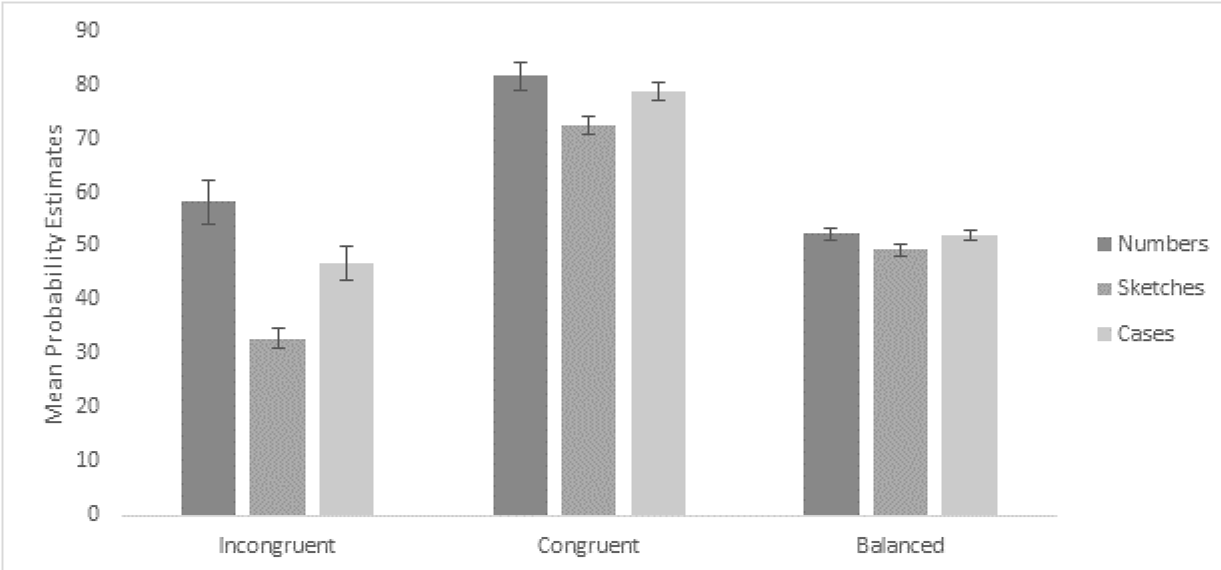

Figure 3.2. Mean probability estimates for Experiment 2 as a function of Strategy x Congruency. Error bars represent standard errors.

3.3.4.2. Confidence.

There was evidence in Experiment 1 that the Sketches strategy group was less sensitive to conflict (as indexed by confidence) than the other strategy groups; participants who reported applying a Sketches strategy were more confident on conflict problems than participants who reported the Numbers and Cases strategies. In Experiment 2, the effect of strategy on confidence was not significant, $F<1$, but there was a Congruency x Strategy interaction, $F(4,324)=3.41$, $p=.01$, $\eta_p^2=0.04$ (see Figure 3.3). To decompose this interaction, a one-way ANOVA was calculated within each congruency condition for the explicit-awareness group factor.

Contrary to the prediction, for incongruent problems, the effect of strategy was not significant, $F<1$; the Numbers group ($M=6.89$, $SD=1.02$), Sketches group ($M=7.02$, $SD=1.01$) and Cases group ($M=6.78$, $SD=1.11$) did not differ on their confidence reports for incongruent problems. In the congruent condition, there also was no effect of strategy, $F(2,162)=1.03$, $p=.361$, $\eta_p^2=0.01$. Confidence reported for congruent problems by the Numbers group ($M=7.14$, *SD*=1.16) did not differ from confidence reported by the Sketches group (*M*=7.0, *SD*=0.95) or the Cases group (*M*=7.25, *SD*=0.94); the Sketches and Cases groups also did not differ on reported confidence for congruent problems. The strategy effect was also not significant in the balanced condition, $F(2,162)=1.16$, $p=.318$, $\eta_p^2=0.01$. For balanced problems, confidence reported by the Numbers group (*M*=6.64, *SD*=0.98) did not differ from the Sketches group (*M*=6.96, *SD*=1.0) or the Cases group (*M*=6.8, *SD*=1.11); the Sketches and Cases groups also did not differ on confidence for balanced problems.

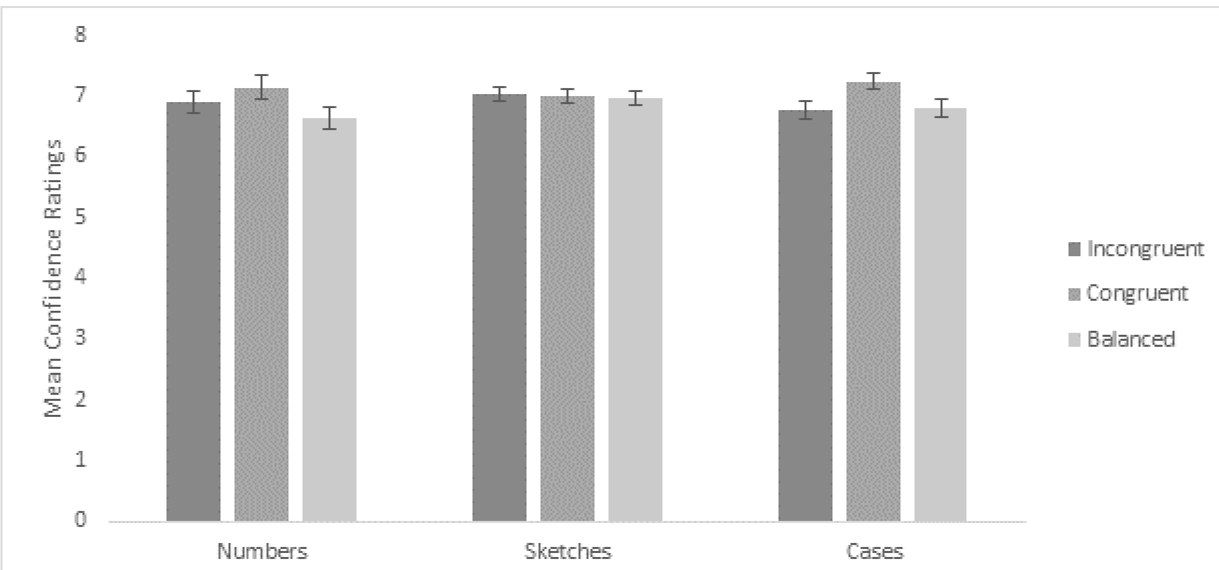

Figure 3.3. Mean confidence ratings for Experiment 2 as a function of Strategy x Congruency. Error bars represent standard errors.

The interaction occurred because the effect of congruency differed between the strategy groups. In the Cases group, the effect of congruency was significant, $F(2,108)=10.01$, $p<.001$, $\eta_p^2=0.16$. Confidence was higher for congruent problems than incongruent problems ($t(54)=4.05$, $p<.001$) and balanced problems ($t(54)=3.49$, $p=.003$), but incongruent and balanced problem confidence did not differ ($t(54)=0.21$, $p>.999$). In the Sketches group, the effect of congruency was not significant, $F<1$. For the Numbers group, the effect of congruency was significant, $F(2,62)=3.45$, $p=.038$, $\eta_p^2=0.1$, but pairwise comparisons found no differences in confidence between any of the congruency conditions ($ts(31)<2.51$, $ps>.052$). Thus, the expected effect of conflict only manifested for those who reported a Cases strategy, suggesting that confidence as an index of conflict detection is sensitive to the resolution strategy adopted by the reasoner.

3.3.4.3. Response time.

It was predicted that the Cases group would take the longest to respond. Instead, response times did not differ across strategy groups, $F(2,162)=1.15$, $p=.319$, $\eta_p^2=0.01$. The Numbers group ($M=12.43$, $SD=2.05$), Sketches group ($M=13.24$, $SD=4.21$), and Cases group ($M=12.2$, $SD=3.36$) did not differ on response times.

3.3.4.4. Explicit Awareness.

In Experiment 1, the Cases group had the highest explicit awareness ratings and the Sketches group had the lowest explicit awareness ratings. However, in Experiment 2, the effect of strategy on explicit awareness was not significant, $F(2,128)=2.32$, $p=.102$, $\eta_p^2=0.04$. Explicit awareness reported by the Numbers group ($M=3.35$, $SD=1.2$), the Sketches group ($M=3.8$, $SD=1.19$), and the Cases group ($M=3.99$, $SD=1.19$) did not differ.

3.4. Discussion

Experiment 2 was designed to influence explicit awareness of conflict by (a) increasing awareness with direct queries to draw the attention of the reasoner to the conflict and (b) decreasing awareness with a concurrent working-memory task. The experimental manipulation was unsuccessful. This result was especially surprising regarding the intervention; alerting reasoners to the conflict had no effect on their awareness. One reasonable interpretation of these data is that influencing explicit awareness of conflict is difficult because reasoners are already highly aware of the conflict. A large majority of participants reported being aware of the conflict and overall, explicit awareness ratings were high in Experiment 2 (replicating Experiment 1).

The replication of the effect of conflict resolution strategy from Experiment 1 was achieved in Experiment 2: conflict resolution strategy was predictive of probability estimates. Additionally, the proportion of reasoners reporting each strategy was consistent in each experiment. These data suggest that base-rate neglect is a function of strategy. For congruent and incongruent problems, reasoners who reported relying on the base-rate ratios to inform their decisions provided probability estimates closest to the base-rate ratios. For incongruent problems, those who reported reliance upon the personality sketches give probability estimates furthest from the base-rate ratios[7]. Additionally, the hallmark conflict detection effect on confidence (lower confidence for incongruent problems relative to congruent problems) was only found for reasoners who reported a Cases strategy, suggesting that sensitivity to conflict (as indexed by confidence) is a function of the conflict resolution strategy adopted by the reasoner to solve the problem.

Chapter 4. General Discussion

The central claim of the conflict detection literature is that detection is processed implicitly and thus, reasoners are relatively unaware of the conflict (De Neys, 2012; 2014). The two experiments in the present thesis provide new and compelling evidence that reasoners demonstrate explicit awareness of conflict and that the primary behavioural measures used to evince detection of conflict are indexing more than the conflict detection process *per se*. In Experiment 1, the patterns of response times, fixation times, and confidence were more

[7] The effect of Strategy was robust in each sample taken for Experiment 2.

consistent with a coherence-resolution interpretation of these indices than a conflict-detection interpretation. Response times were longer for neutral problems than conflict and non-conflict problems, suggesting that response time as an index is sensitive to other processes than conflict detection (i.e., the resolution of conflict). Fixation times were not sensitive to conflict; there were no differences between the conflict and non-conflict problems regarding fixation times. Instead, fixation times were longest for neutral problems on both the base-rate ratios and personality descriptions. The extra time spent on neutral problems appears to be spent evaluating both the base-rate ratios and personality descriptions. In other words, contrary to the conflict detection interpretation, response times and fixation times are most sensitive to the neutral problems, not the conflict in the incongruent problems.

Confidence ratings were more consistent with an effect of coherence monitoring (Koriat, 2012) than a conflict-detection interpretation (De Neys, 2014). Under the conflict-detection interpretation, confidence would be expected to be rated lowest for the problems with the highest degree of conflict. Two results contradicted this prediction. First, confidence was lowest for neutral problems, not the conflict problems. Second, confidence was not sensitive to the extremity of the base-rate ratios (i.e., the degree of conflict between ratios and descriptions) for conflict problems but was sensitive to the extremity of the ratios for the non-conflict congruent problems. These data are difficult to reconcile with the conflict-detection interpretation. The relative coherence of the information presented to the reasoner appears to better account for these confidence data than the presence of conflicting responses to the problem.

Across both experiments, self-reported resolution strategies were robustly associated with probability estimates. These data suggest that base-rate neglect is a function of the strategic approach to resolve the conflict and the measures intended to index detection of conflict are also indexing conflict resolution strategy. Also, in two experiments, reasoners reported high explicit awareness of conflict. There is reason to trust these self-reports. Reasoner's retrospective reports on their resolution strategy mapped sensibly on to a host of behavioural measures: probability estimates, confidence ratings, response times, and fixation times. It is reasonable to suggest that the lack of correlation between the so-called implicit measures of conflict detection (confidence ratings, response times, and fixation times) and the novel self-reports of explicit conflict awareness indicates that they are indexing different cognitive factors, namely, measures of the explicit awareness of conflict and measures that index multiple processes: coherence monitoring, conflict detection, and conflict resolution.

4.1. Base-Rate Neglect and Conflict Resolution Strategy

The typical base-rate neglect phenomenon was replicated in both experiments: probability estimates were further from the base-rate probability for conflict problems than non-conflict problems (De Neys & Glumicic, 2008; Newman et al., 2017; Pennycook & Thompson, 2012; Thompson et al., 2011). The pattern of probability estimates for neutral personality descriptions (Newman et al., 2017) was also replicated: estimates for neutral problems were closer to the base-rate probability and further from the base-rate probability than estimates for conflict and non-conflict problems, respectively.

The base-rate neglect phenomenon in these experiments appeared to be a function of conflict resolution strategy. In both experiments, a robust effect of the self-reported strategy was found: reasoners who reported relying on the base-rate ratios gave responses closest to the base-rate probability and reasoners who reported relying on the personality descriptions gave responses furthest from the base-rate probability. Moreover, a consistent proportion of reasoners reported adopting each of the strategy options in both experiments: approximately 20% of reasoners relied on the ratios, 45% relied on the descriptions, and 35% took a case-by-case approach. This distribution of strategies is consistent with the position that base-rate neglect occurs because reasoners who resolve conflict in favour of the personality descriptions outnumber reasoners who resolve conflict in favour of the base-rate probability (Newman et al., 2017; Pennycook & Thompson, 2012). In sum, these data suggest that the base-rate task is fundamentally a task that assesses conflict resolution preferences, not a uniform inability to recognize and avoid a compelling lure (i.e., the stereotypical personality description).

4.2. Behavioural Indices are Sensitive to Multiple Factors

The typical effects of conflict were replicated in two experiments: reasoners were less confident and took longer to respond when solving conflict problems relative to non-conflict congruent problems. A crucial test of the conflict-detection and coherence-resolution interpretations were the response times, fixation times, and confidence ratings for neutral problems in Experiment 1. Neutral problems are also non-conflict problems. Therefore, the conflict-detection and coherence-resolution interpretations made different predictions on how the presumed indices of conflict detection would manifest for the neutral problems relative to the conflict problems. For each of the measures (response times, fixation times, and confidence), the evidence was clear: the hallmark indicators of the detection of conflict (reduced confidence, longer response times, longer fixation times) were even more pronounced for neutral problems than for incongruent conflict problems (see Figures 2.5, 2.6, 2.7, and 2.8).

The predominant interpretation is that differences between congruent and incongruent problems are an effect of conflict detection (De Neys, 2014). If this is the case, then how can a larger conflict detection effect occur between congruent and neutral problems than congruent and incongruent problems? The neutral problem data is inconsistent with the conflict-detection interpretation. From these data, we can conclude that these measures index something else beyond conflict detection. Possibilities include the resolution of conflict, indecision or uncertainty, or relative coherence of the information. For example, monitoring of coherence can influence fluency (Koriat, 2012), which in turn impacts response times and confidence (Topolinski, 2011). Another possibility is that uninformative neutral problems are interpreted as especially odd by reasoners, effectively slowing down their reasoning and increasing their uncertainty. These indices are likely sensitive to multiple cognitive factors, some of which can also elicit behavioural responses that mirror the presumed effects of conflict detection. If a measure indexes multiple cognitive factors, it is no longer warranted to infer from the effect (e.g., reduced confidence for conflict problems) to the cause (e.g., conflict detection) because there are multiple potential causes. Considering the present evidence, the logic of the conflict-detection interpretation is less sound.

The difference on behavioural indices between conflict and non-conflict problems may represent two separate dimensions: a conflict effect for incongruent problems and a coherence effect for congruent problems (Koriat, 2012). If this is the case, then when a difference is recorded between conflict and non-conflict problems, how can one discriminate whether the observed effect was present in one condition, the other condition, or both conditions?

In Experiment 1, there was evidence that the typical "conflict" effect on confidence was a decrease in confidence for conflict problems. Rather than attributing the difference to the presence of conflict, it might have something to do with the coherence of the information. There was no effect of extremity for the incongruent problems but there was an effect of extremity for the congruent problems; this is inconsistent with the conflict-detection interpretation. This evidence suggests that the conflict effect is not due to relatively lower confidence in the presence of conflict, but rather, relatively higher confidence in the presence of coherent information. Moreover, the conflict effect did not differ as a function of the degree of conflict between the base-rate ratio and personality description. Instead, it was observed that confidence increased for congruent problems as a function of the consistency of the evidence (i.e., the interaction of congruency and extremity).

These data suggest that much of the evidence in support of the conflict-detection interpretation (De Neys, 2014) hinges upon behavioural measures that are sensitive to more than just detection of conflict. Previously, much of the conflict detection evidence has been lumped together as a long list of behavioural indicators (such as response times, confidence, fixation times, and others) that differentiate congruent and incongruent problems (see De Neys, 2014). Implicit in this claim is the assumption that these measures all index the same underlying cognitive process. Our data challenge this assumption. Instead, these measures should be discriminated based on the aspects of conflict detection, resolution, and coherence monitoring processes that they index. This research is a first step towards determining where these measures converge and where they diverge.

4.3. Explicit Awareness of Conflict

One of the primary hypotheses tested in this research was whether reasoners are explicitly aware of conflict. The prediction was that reasoners are more aware than previously assumed by the conflict-detection interpretation. There was evidence to support that claim: reasoners reported high awareness of conflict (on a scale from 1 to 7) in both Experiment 1 (M=4.14, SD=0.98) and Experiment 2 (M=3.79, SD=1.2). Additionally, a large majority of reasoners in Experiment 2 responded "yes" to a yes-no question of their awareness of competing responses (approximately 80%). Therefore, reasoners appeared to be relatively aware of the conflict, explicitly.

The retrospective strategy reports taken were phrased to determine the information a reasoner relies upon to reach their probability estimates. In other words, it is a question of how one resolved the conflict, not how one may or may not have identified the conflict. The strategy reports were robustly associated with a variety of behavioural measures that supported their validity. Therefore, this measure was an accurate assessment of conflict resolution strategy.

The strategy reports demonstrated how resolution strategy affected the supposed implicit measures of conflict detection. In Experiment 1, reasoners who applied a Sketches strategy reported higher confidence on conflict problems than those who reported a Numbers or Cases strategy. The Sketches strategy group may be confident when they can resolve the conflict in favour of their preferred source of information: the personality descriptions. Notably, this effect was absent in Experiment 2; perhaps this effect of strategy is contextual, dependent on the presence of neutral problems in the task. Additionally, the reasoners who reported a Cases strategy displayed longer response times than the other strategies. It may be that the strategy that weights the probabilistic and stereotypical information more equally required a longer time to resolve the conflict between said information.

Clearly, these measures are sensitive to how a reasoner resolves the conflict in the problems. Even if these measures were sensitive to an implicit process of conflict detection, they are also indexing the resolution of that conflict. Thus, when the hallmark indicators of supposed conflict detection are observed (e.g., reduced confidence, longer response time), it is unclear what the underlying cognitive processes were to elicit those behaviours.

Ultimately, measures that index conflict resolution processes would likely not be correlated with awareness of the conflict itself. Explicit awareness of conflict is a measure of how aware a reasoner was that there was some degree of inconsistency in the problem or that there were competing responses available. Measures such as confidence and response times appear to be sensitive to what the reasoner does once they are aware of the conflict. Additionally, there are clear individual differences in resolution strategy: the distribution of strategies reported is not uniform and reported strategy is predictive of performance on the base-rate task. Thus, it is not surprising that awareness was not correlated with the typical measures of conflict detection.

The attempt to increase and decrease explicit awareness of conflict in Experiment 2 failed. In retrospect, there is a straightforward explanation for this outcome: reasoners were already aware of the conflict. If so, it would be difficult to either prevent or facilitate detection of the conflict. These data suggest that reasoners are explicitly aware of the conflict during problem solving, but differ considerably on how they react to and resolve the conflict. Moving forward, efforts to develop more nuanced measures are required. Identifying new paradigms that allow for per-trial measures of explicit awareness and resolution strategy are needed.

4.4. Conclusion

The evidence gathered here provided a strong challenge to the prevailing conflict-detection interpretation, suggesting that reasoners are explicitly aware of conflict but they differ on their resolution of said conflict. The data are consistent with the conclusion that the typical indices of conflict detection are sensitive to conflict resolution and coherence monitoring processes. The conflict detection literature implies a flawed reasoning process: detecting conflict is an indicator that something is amiss in the information presented and failing to resolve it in favour of the statistically or logically correct response is an error. I have argued here for a more positive position: it is the subjective evaluation of the evidence to reach a self-consistent conclusion that drives what have been considered "conflict" effects. Current reasoning theories need to move from the poorly specified conflict detection stage toward a broader view that can account for the advantageous recognition of coherent information and resolution of uncertainty.

References

Ackerman, R., & Thompson, V. (2017). Meta-Reasoning: Monitoring and Control of Thinking and Reasoning, (June). https://doi.org/10.1016/j.tics.2017.05.004

Ariel, R., & Castel, A. D. (2014). Eyes wide open: Enhanced pupil dilation when selectively studying important information. *Experimental Brain Research*, *232*(1), 337–344. https://doi.org/10.1007/s00221-013-3744-5

Ball, L. J., Phillips, P., Wade, C. N., & Quayle, J. D. (2006). Effects of belief and logic on syllogistic reasoning: Eye-movement evidence for selective processing models. *Experimental Psychology*, *53*(1), 77–86. https://doi.org/10.1027/1618-3169.53.1.77

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, *30*(3), 241–297. https://doi.org/10.1017/S0140525X07001653

Bethell-Fox, C. E., & Shepard, R. N. (1988). Mental rotation: Effects of stimulus complexity and familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(1), 12–23. https://doi.org/10.1037/0096-1523.14.1.12

Blignaut, P. (2009). Fixation identification: The optimum threshold for a dispersion algorithm. *Attention, Perception & Psychophysics*, *71*(4), 881–895. https://doi.org/10.3758/APP

Bonner, C., & Newell, B. R. (2010). In conflict with ourselves? An investigation of heuristic and analytic processes in decision making. *Memory & Cognition*, *38*(2), 186–196. https://doi.org/10.3758/MC.38.2.186

Cohen, A. D. (1987). Using verbal reports in research on language learning. In C. Faerch & G. Kasper (Eds.), Introspection in second language research (pp. 82-95). Philadelphia: Multilingual Matters, Ltd.

De Neys, W. (2006). Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual-task considerations. *Quarterly Journal of Experimental Psychology (2006)*, *59*(6), 1070–1100. https://doi.org/10.1080/02724980543000123

De Neys, W. (2012). Bias and Conflict: A Case for Logical Intuitions. *Perspectives on Psychological Science*, *7*(1), 28–38. https://doi.org/10.1177/1745691611429354

De Neys, W. (2014). Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning*, *20*(2), 169–187. https://doi.org/10.1080/13546783.2013.854725

De Neys, W., Cromheeke, S., & Osman, M. (2011). Biased but in doubt: Conflict and decision confidence. *PLoS ONE*, *6*(1). https://doi.org/10.1371/journal.pone.0015954

De Neys, W., & Feremans, V. (2012). Development of Heuristic Bias Detection in Elementary School. *Developmental Psychology*, *49*(2), 258–269. https://doi.org/10.1037/a0028320

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*(3), 1248–1299. https://doi.org/10.1016/j.cognition.2007.06.002

De Neys, W., Moyens, E., & Vansteenwegen, D. (2010). Feeling we're biased: autonomic arousal and reasoning conflict. *Cognitive, Affective & Behavioral Neuroscience*, *10*(2), 208–216. https://doi.org/10.3758/CABN.10.2.208

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition : Advancing the Debate. https://doi.org/10.1177/1745691612460685

Ericsson, K. A., & Simon, H. A. (1987). Verbal reports on thinking. In C. Faerch & G. Kasper (Eds.), Introspection in second language research (pp. 24-53). Philadelphia: Multilingual Matters, Ltd.

Franssens, S., & De Neys, W. (2009). The effortless nature of conflict detection during thinking. *Thinking & Reasoning*, *15*(2), 105–128. https://doi.org/10.1080/13546780802711185

Handley, S. J., & Trippas, D. (2015). Dual processes and the interplay between knowledge and structure: A new parallel processing model. *Psychology of Learning and Motivation - Advances in Research and Theory*, *62*(September), 33–58. https://doi.org/10.1016/bs.plm.2014.09.002

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction: Whose is the fallacy? *Cognition*, *7*(4), 385–407. https://doi.org/10.1016/0010-0277(79)90023-4

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review*, *119*(1), 80–113. https://doi.org/10.1037/a0025648

Manor, B. R., & Gordon, E. (2003). Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. *Journal of Neuroscience Methods*, *128*(1–2), 85–93. https://doi.org/10.1016/S0165-0270(03)00151-1

Mele, M. L., & Federici, S. (2012). Gaze and eye-tracking solutions for psychological research. *Cognitive Processing*, *13*(1 SUPPL). https://doi.org/10.1007/s10339-012-0499-z

Mevel, K., Poirel, N., Rossi, S., Cassotti, M., Simon, G., Houdé, O., & De Neys, W. (2015). Bias detection: Response confidence evidence for conflict sensitivity in the ratio bias task. *Journal of Cognitive Psychology*, *27*(2), 227–237. https://doi.org/10.1080/20445911.2014.986487

Morsanyi, K., & Handley, S. J. (2012). Logic feels so good—I like it! Evidence for intuitive detection of logicality in syllogistic reasoning. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, *38*(3), 596–616. https://doi.org/10.1037/a0026099

Naber, M., Frässle, S., Rutishauser, U., & Einhäuser, W. (2013). Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. *Journal of Vision*, *13*(2), 11. https://doi.org/10.1167/13.2.11.doi

Newman, I. R., Gibb, M., & Thompson, V. A. (2017). Rule-Based Reasoning Is Fast and Belief-Based Reasoning Can Be Slow : Challenging Current Explanations of Belief-Bias and Base-Rate Neglect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(1). https://doi.org/10.1037/xlm0000372

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2012). Are we good at detecting conflict during reasoning? *Cognition*, *124*(1), 101–106. https://doi.org/10.1016/j.cognition.2012.04.004

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, *80*, 34–72. https://doi.org/10.1016/j.cogpsych.2015.05.001

Pennycook, G., & Thompson, V. A. (2012). Reasoning with base rates is routine, relatively effortless, and context dependent. *Psychonomic Bulletin & Review*, *19*(3), 528–534. https://doi.org/10.3758/s13423-012-0249-3

Pennycook, G., Trippas, D., Handley, S. J., & Thompson, V. A. (2014). Base Rates : Both Neglected and Intuitive, *40*(2), 544–554. https://doi.org/10.1037/a0034887

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from http://www.pstnet.com

Russo, J., Johnson, E., & Stephens, D. (1989). The validity of verbal protocols. *Memory & Cognition*, *17*(6), 759–769. https://doi.org/10.3758/BF03202637

SensoMotoric Instruments, Inc. [BeGaze 3.6]. (2016). http://www.smivision.com

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: implications for the rationality debate? *The Behavioral and Brain Sciences*, *23*(5), 645-665-726. https://doi.org/10.1017/S0140525X00003435

Taylor, K. L., & Dionne, J.-P. (2000). Accessing problem-solving strategy knowledge: The complementary use of concurrent verbal protocols and retrospective debriefing. *Journal of Educational Psychology*, *92*(3), 413–425. https://doi.org/10.1037/0022-0663.92.3.413

Thompson, V. A., & Johnson, S. C. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning*, *20*(January 2015), 215–244. https://doi.org/10.1080/13546783.2013.869763

Thompson, V. A., Prowse-Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107–140. https://doi.org/10.1016/j.cogpsych.2011.06.001

Thompson, V. A., Turner, J. A. P., Pennycook, G., Ball, L. J., Brack, H., Ophir, Y., & Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition*, *128*(2), 237–251. https://doi.org/10.1016/j.cognition.2012.09.012

Toplak, M., West, R., & Stanovich, K. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, *39*(7), 1275–1289. https://doi.org/10.3758/s13421-011-0104-1

Topolinski, S. (2011). A process model of intuition. *European Review of Social Psychology*, *22*(1), 274–315. https://doi.org/10.1080/10463283.2011.640078

Trbovich, P. L., & LeFevre, J.-A. (2003). Phonological and visual working memory in mental addition. *Memory & Cognition*, *31*(5), 738–745. https://doi.org/10.3758/BF03196112

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty : Heuristics and Biases, *185*(4157), 1124–1131.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology*: I (pp. 106–137). Harmandsworth, UK: Penguin.

Wilson, T. D. (1994). Commentary to Feature Review Validity and Completeness of Verbal Reports, *5*(5), 249–252.

Appendices

Appendix A

To change an incongruent problem to congruent (or vice versa), reverse the values in the base-rate. For example, the incongruent problem Q1H has 5 engineers and 995 lawyers and the description of Jack fits the stereotype of an engineer (incongruent). To change the problem to congruent, change the base-rate ratio to 995 engineers and 5 lawyers. Neutral problems are neither congruent nor incongruent.

<u>Incongruent Problems</u>

In a study 1000 people were tested. Among the participants there were 5 engineers and 995 lawyers. Jack is a randomly chosen participant of this study. Jack is 36 years old. He is not married and is somewhat introverted. He likes to spend his free time reading science fiction and writing computer programs. What is the probability that Jack is a lawyer?

In a study 1000 people were tested. Among the participants there were 300 men and 700 women. Jessie is a randomly chosen participant of this study. Jessie is 23 years old and is finishing a degree in engineering. On Friday nights, Jessie likes to go out cruising with friends while listening to loud music and drinking beer. What is the probability that Jessie is a man?

In a study 1000 people were tested. Among the participants there were 470 who live in a condo and 530 who live in a farmhouse. Kurt is a randomly chosen participant of this study. Kurt works on Wall Street and is single. He works long hours and wears Armani suits to work. He likes wearing shades. What is the probability that Kurt lives in a farmhouse?

In a study 1000 people were tested. Among the participants there were 996 nurses and 4 doctors. Paul is a randomly chosen participant of this study. Paul is 34 years old. He lives in a beautiful home in a posh suburb. He is well spoken and very interested in politics. He invests a lot of time in his career. What is the probability that Paul is a doctor?

In a study 1000 people were tested. Among the participants there were 290 whose favorite series is Star Trek and 710 whose favorite series is Days of Our Lives. Jeremy is a randomly chosen participant of this study. Jeremy is 26 and is doing graduate studies in physics. He stays at home most of the time and likes to play video-games. What is the probability that Jeremy's favorite series is Days of Our Lives?

In a study 1000 people were tested. Among the participants there were 480 sixteen- year olds and 520 fifty-year olds. Ellen is a randomly chosen participant of this study. Ellen likes to listen to hip hop and rap music. She enjoys wearing tight shirts and jeans. She's fond of dancing and has a small nose piercing. What is the probability that Ellen is sixteen?

In a study 1000 people were tested. Among the participants there were 3 accountants and 997 street artists. Brannon is a randomly chosen participant of this study. Brannon is 29 years old. He is very good with numbers but is shy around people. He spends much of his time working. What is the probability that Brannon is an accountant?

In a study 1000 people were tested. Among the participants there were 280 writers and 720 construction workers. Hank is a randomly chosen participant of this study. Hank is 42 years old. He is a creative and introverted person. He considers his home computer his most prized possession. What is the probability that Hank is a writer?

In a study 1000 people were tested. Among the participants there were 510 managers and 490 firemen. Tyrone is a randomly chosen participant of this study. Tyrone is 27 years old. All his friends consider him very brave and he is in relatively good physical shape. He goes to the gym regularly. What is the probability that Tyrone is a manager?

Congruent Problems

In a study 1000 people were tested. Among the participants there were 995 who buy their clothes at high-end retailers and 5 who buy their clothes at Wal-Mart. Karen is a randomly chosen participant of this study. Karen is a 33-year-old female. She works in a business office and drives a Porsche. She lives in a fancy penthouse with her boyfriend. What is the probability that Karen buys her clothes at Wal-Mart?

In a study 1000 people were tested. Among the participants there were 700 girls and 300 boys. Kelly is a randomly chosen participant of this study. Kelly is 13 years old. Kelly's favourite subject is art. Kelly's favourite things to do are shopping and having sleepovers with friends to gossip about other kids at school. What is the probability that Kelly is a boy?

In a study 1000 people were tested. Among the participants there were 530 who have a tattoo and 470 without tattoo. Jay is a randomly chosen participant of this study. Jay is a 29-year-old male. He has served a short time in prison. He has been living on his own for 2 years now. He has an older car and listens to punk music. What is the probability that Jay has a tattoo?

In a study 1000 people were tested. Among the participants there were 996 kindergarten teachers and 4 executive managers. Lilly is a randomly chosen participant of this study. Lilly is 37 years old. She is married and has 3 kids. Her husband is a veterinarian. She is committed to her family and always watches the daily cartoon shows with her kids. What is the probability that Lilly is a kindergarten teacher?

In a study 1000 people were tested. Among the participants there were 290 Bruce Springsteen fans and 710 Britney Spears fans. Tara is a randomly chosen participant of this study. Tara is 15. She loves to go shopping at the mall and to talk with her friends about their crushes at school. What is the probability that Tara is a Bruce Springsteen fan?

In a study 1000 people were tested. Among the participants there were 480 Americans and 520 French people. Martine is a randomly chosen participant of this study. Martine is 26 years old. She is bilingual and reads a lot in her spare time. She is a very fashionable dresser and a great cook. What is the probability that Martine is French?

In a study 1000 people were tested. Among the participants there were 997 aeroplane pilots and 3 shop assistants. George is a randomly chosen participant of this study. George is 36 years old.

He is very intelligent and has nerves of steel. He has great hand-eye coordination. What is the probability that George is a shop assistant?

In a study 1000 people were tested. Among the participants there were 280 I.T. Technicians and 720 politicians. Richard is a randomly chosen participant of this study. Richard is 38 years old. He is a good public speaker and is good at meeting people. He is a top notch debater and can argue both sides of an issue with ease. What is the probability that Richard is a politician?

In a study 1000 people were tested. Among the participants there were 510 paramedics and 490 clowns. Dan is a randomly chosen participant of this study. Dan is 30 years old. He is a good driver and a takes his job very seriously. He is married, but has no children. What is the probability that Dan is a clown?

Neutral Problems

In a study 1000 people were tested. Among the participants there were 5 who campaigned for George W. Bush and 995 who campaigned for John Kerry. Jim is a randomly chosen participant of this study. Jim is 5 ft and 8 in. tall, has black hair, and is the father of two young girls. He drives a yellow van that is completely covered with campaign posters. What is the probability that Jim campaigned for John Kerry?

In a study 1000 people were tested. Among the participants there were 700 men and 300 women. Casey is a randomly chosen participant of this study. Casey is a 36-year-old writer. Casey has two brothers and one sister. Casey likes running and watching a good movie. What is the probability that Casey is a woman?

In a study 1000 people were tested. Among the participants there were 470 who play the drums and 530 who play the saxophone. Tom is a randomly chosen participant of this study. Tom is 20 years old. He is studying in Washington and has no steady girlfriend. He just bought a second-hand car with his savings. What is the probability that Tom plays the drums?

In a study 1000 people were tested. Among the participants there were 996 pool players and 4 basketball players. Jason is a randomly chosen participant of this study. Jason is 29 years old and has lived his whole life in New York. He has green colored eyes and black hair. He drives a light-gray colored car. What is the probability that Jason is a basketball player?

In a study 1000 people were tested. Among the participants there were 290 who live in New York and 710 who live in Los Angeles. Christopher is a randomly chosen participant of this study. Christopher is 28 years old. He has a girlfriend and shares an apartment with a friend. He likes watching basketball. What is the probability that Christopher lives in Los Angeles?

In a study 1000 people were tested. Among the participants there were 480 computer science majors and 520 English majors. Matt is a randomly chosen participant of this study. Matt is 20 years old and lives in downtown Toronto. Matt’s favourite food is pasta with meatballs. His parents are living in Vancouver. What is the probability that Matt is a computer science major?

In a study 1000 people were tested. Among the participants there were 997 blondes and 3 brunettes. Geraldine is a randomly chosen participant of this study. Geraldine is 41 years old. She loves a good movie and spends a lot of her free time reading. She enjoys helping her two children with their homework. What is the probability that Geraldine is a brunette?

In a study 1000 people were tested. Among the participants there were 720 bookkeepers and 280 bank tellers. Dianna is a randomly chosen participant of this study. Dianna is 59 years old. She has been employed at her current job for 7 years. She recently took a holiday to visit with family and friends. What is the probability that Dianna is a bookkeeper?

In a study 1000 people were tested. Among the participants there were 510 teachers and 490 secretaries. Molly is a randomly chosen participant of this study. Molly is 25 years old. She drives a mid-sized car and lunches in the cafeteria twice a week. She enjoys hiking and dancing. What is the probability that Molly is a secretary?

Appendix B

Q1

Which of the following describes how you answered the questions:

1. I mostly relied on the numbers
2. I mostly paid attention to the personality sketches
3. I guessed randomly
4. I evaluated each problem on a case-by-case basis

   If so, what was your criterion? Enter in the text box below

5. None of the above.

   Please describe in the text box below

Please enter a number 1-5 to select your response.

If you select 4 or 5, please enter additional text in the box provided below.

Press TAB on the keyboard to complete the question.

Q2

Did you feel that there was two (or more) answers, which were quite different from one another, to some of the problems?

1 2 3 4 5 6 7

Never Sometimes Frequently Always

Q3

Was the following occupation described in the problems that you just solved?

[Occupation]

Yes - Y No - N

-librarian

-doctor

-professor

-accountant

-cashier

-politician

Q4

Please estimate how long it took you to complete all 27 problems, from start to finish, then press ENTER:

Q5

Did you notice that sometimes the personality sketch and the numbers suggested different responses?

1 2 3 4 5 6 7

Never Sometimes Frequently Always

Q6
Was the following value amongst the sample sizes described in the problems that you just solved?
[Sample]
Yes - Y No - N
-900
-330
-490
-700
-997
-500

Q7
Did you ever feel torn between two choices and needed to decide which you felt was a better solution?
1 2 3 4 5 6 7
Never Sometimes Frequently Always

Q8
How did you answer the questions when you noticed that the personality sketch and the numbers suggested different responses or where you felt torn between two choices?
Please enter your response in the text box below, and press TAB to finish.

Appendix C

In a big research project a number of studies were carried out where short personality descriptions of the participants were made. In every study there were participants from two population groups (e.g., carpenters and policemen). In each study one participant was drawn at random from the sample. You will see the personality description of this randomly chosen participant. You will also get information about the composition of the population groups tested in the study of question.

For each problem you'll be asked to provide us with a probability judgment out of 100, indicating what you believe the likelihood is that the randomly chosen participant belongs to the specified group.

For example, if you are sure that the participant belongs to the specified group, you should choose a number closer to 100. Conversely, if you are sure that the participant does not belong to the specified group, you should choose a number closer to 0. Finally, if you are unsure about which group the participant belongs to, then you should choose a number close to 50.

For all problems, the group that you are asked to give the probability for was randomly chosen out of the two groups presented.

For each problem, we want you to provide two answers. For the first answer, we are interested in the answer that is your FIRST INCLINATION or INSTINCT. Thus, as soon as you have read the problem, please provide us with the FIRST ANSWER that comes to mind. It is important that you try to answer as quickly as possible, as you will only have 12 seconds in which to provide your answer. Once 12 seconds has elapsed the problem will become italicized and change color. At this point you must enter your answer IMMEDIATELY if you have not yet done so. For this answer you will not be allowed to change your response once given.

Your second answer should be your FINAL ANSWER. For this answer you will have as much time as you like to answer. Please make sure that you take your time and think about the problem carefully. At this point, you will be able to make changes to your answer if necessary.

After each response, you will be asked to rate how confident you are with your response. For your first answer this scale will measure how ‘right' you feel about the answer. 1 will correspond with "doesn't feel right at all" and 9 will correspond with "feels very right".

For your second answer, this scale will measure how confident you are with your final answer. 1 will correspond with "not at all confident" and 9 will correspond with "extremely confident".

Once you have provided a response, you will have a couple of seconds before the next problem will appear on the screen.

Please feel free to ask any questions that you have.

Appendix D

To change an incongruent problem to congruent (or vice versa), reverse the values in the base-rate. For example, the first incongruent problem has 5 engineers and 995 lawyers and the description of Jack fits the stereotype of an engineer (incongruent). To change the problem to congruent, change the base-rate ratio to 995 engineers and 5 lawyers. To change any of the problems to balanced problems, change the base-rate ratio to values close to 50:50 (e.g., 520:480).

Problems

In a study 1000 people were tested. Among the participants there were 5 engineers and 995 lawyers. Jack is a randomly chosen participant of this study. Jack is 36 years old. He is not married and is somewhat introverted. He likes to spend his free time reading science fiction and writing computer programs. What is the probability that Jack is a lawyer?

In a study 1000 people were tested. Among the participants there were 300 men and 700 women. Jessie is a randomly chosen participant of this study. Jessie is 23 years old and is finishing a degree in engineering. On Friday nights, Jessie likes to go out cruising with friends while listening to loud music and drinking beer. What is the probability that Jessie is a man?

In a study 1000 people were tested. Among the participants there were 470 who live in a condo and 530 who live in a farmhouse. Kurt is a randomly chosen participant of this study. Kurt works on Wall Street and is single. He works long hours and wears Armani suits to work. He likes wearing shades. What is the probability that Kurt lives in a farmhouse?

In a study 1000 people were tested. Among the participants there were 996 nurses and 4 doctors. Paul is a randomly chosen participant of this study. Paul is 34 years old. He lives in a beautiful home in a posh suburb. He is well spoken and very interested in politics. He invests a lot of time in his career. What is the probability that Paul is a doctor?

In a study 1000 people were tested. Among the participants there were 290 whose favorite series is Star Trek and 710 whose favorite series is Days of Our Lives. Jeremy is a randomly chosen participant of this study. Jeremy is 26 and is doing graduate studies in physics. He stays at home most of the time and likes to play video-games. What is the probability that Jeremy's favorite series is Days of Our Lives?

In a study 1000 people were tested. Among the participants there were 480 sixteen- year olds and 520 fifty-year olds. Ellen is a randomly chosen participant of this study. Ellen likes to listen to hip hop and rap music. She enjoys wearing tight shirts and jeans. She's fond of dancing and has a small nose piercing. What is the probability that Ellen is sixteen?

In a study 1000 people were tested. Among the participants there were 3 accountants and 997 street artists. Brannon is a randomly chosen participant of this study. Brannon is 29 years old. He is very good with numbers but is shy around people. He spends much of his time working. What is the probability that Brannon is an accountant?

In a study 1000 people were tested. Among the participants there were 280 writers and 720 construction workers. Hank is a randomly chosen participant of this study. Hank is 42 years old. He is a creative and introverted person. He considers his home computer his most prized possession. What is the probability that Hank is a writer?

In a study 1000 people were tested. Among the participants there were 510 managers and 490 firemen. Tyrone is a randomly chosen participant of this study. Tyrone is 27 years old. All his friends consider him very brave and he is in relatively good physical shape. He goes to the gym regularly. What is the probability that Tyrone is a manager?

In a study 1000 people were tested. Among the participants there were 995 who buy their clothes at high-end retailers and 5 who buy their clothes at Wal-Mart. Karen is a randomly chosen participant of this study. Karen is a 33-year-old female. She works in a business office and drives a Porsche. She lives in a fancy penthouse with her boyfriend. What is the probability that Karen buys her clothes at Wal-Mart?

In a study 1000 people were tested. Among the participants there were 700 girls and 300 boys. Kelly is a randomly chosen participant of this study. Kelly is 13 years old. Kelly's favourite subject is art. Kelly's favourite things to do are shopping and having sleepovers with friends to gossip about other kids at school. What is the probability that Kelly is a boy?

In a study 1000 people were tested. Among the participants there were 530 who have a tattoo and 470 without tattoo. Jay is a randomly chosen participant of this study. Jay is a 29-year-old male. He has served a short time in prison. He has been living on his own for 2 years now. He has an older car and listens to punk music. What is the probability that Jay has a tattoo?

In a study 1000 people were tested. Among the participants there were 996 kindergarten teachers and 4 executive managers. Lilly is a randomly chosen participant of this study. Lilly is 37 years old. She is married and has 3 kids. Her husband is a veterinarian. She is committed to her family and always watches the daily cartoon shows with her kids. What is the probability that Lilly is a kindergarten teacher?

In a study 1000 people were tested. Among the participants there were 290 Bruce Springsteen fans and 710 Britney Spears fans. Tara is a randomly chosen participant of this study. Tara is 15. She loves to go shopping at the mall and to talk with her friends about their crushes at school. What is the probability that Tara is a Bruce Springsteen fan?

In a study 1000 people were tested. Among the participants there were 480 Americans and 520 French people. Martine is a randomly chosen participant of this study. Martine is 26 years old. She is bilingual and reads a lot in her spare time. She is a very fashionable dresser and a great cook. What is the probability that Martine is French?

In a study 1000 people were tested. Among the participants there were 997 aeroplane pilots and 3 shop assistants. George is a randomly chosen participant of this study. George is 36 years old. He is very intelligent and has nerves of steel. He has great hand-eye coordination. What is the probability that George is a shop assistant?

In a study 1000 people were tested. Among the participants there were 280 I.T. Technicians and 720 politicians. Richard is a randomly chosen participant of this study. Richard is 38 years old. He is a good public speaker and is good at meeting people. He is a top notch debater and can argue both sides of an issue with ease. What is the probability that Richard is a politician?

In a study 1000 people were tested. Among the participants there were 510 paramedics and 490 clowns. Dan is a randomly chosen participant of this study. Dan is 30 years old. He is a good driver and a takes his job very seriously. He is married, but has no children. What is the probability that Dan is a clown?

In a study 1000 people were tested. Among the participants there were 997 blondes and 3 brunettes. Geraldine is a randomly chosen participant of this study. Geraldine is 41 years old. She loves a good movie and spends a lot of her free time reading. She enjoys helping her two children with their homework. What is the probability that Geraldine is a brunette?

In a study 1000 people were tested. Among the participants there were 720 bookkeepers and 280 bank tellers. Dianna is a randomly chosen participant of this study. Dianna is 59 years old. She has been employed at her current job for 7 years. She recently took a holiday to visit with family and friends. What is the probability that Dianna is a bookkeeper?

In a study 1000 people were tested. Among the participants there were 510 teachers and 490 secretaries. Molly is a randomly chosen participant of this study. Molly is 25 years old. She drives a mid-sized car and lunches in the cafeteria twice a week. She enjoys hiking and dancing. What is the probability that Molly is a secretary?

In a study 1000 people were tested. Lucius is a randomly chosen participant of this study. Among the participants there were 530 hippies and 470 boxers. Lucius is 34 years old. He is pretty aggressive and tends to get involved in bar fights more than the average person. He recently got divorced. What is the probability that Lucius is a boxer?

In a study 1000 people were tested. Floyd is a randomly chosen participant of this study. Among the participants there were 997 artists and 3 consultants. Floyd 40 years old. He is an imaginative person and enjoys street theatre. He loves experimenting with different types of food. What is the probability that Floyd is an artist?

In a study 1000 people were tested. Corinne is a randomly chosen participant of this study. Among the participants there were 490 gardeners and 510 secretaries. Corinne is 32 years old. She is a great organizer and always dresses neatly. She loves talking to her friends and family on the phone. What is the probability that Corinne is a secretary?