



ORBIT - Online Repository of Birkbeck Institutional Theses

Enabling Open Access to Birkbeck's Research Degree output

Rationality, pragmatics, and sources

<https://eprints.bbk.ac.uk/id/eprint/40284/>

Version: Full Version

Citation: Collins, Peter J. (2017) Rationality, pragmatics, and sources. [Thesis] (Unpublished)

© 2020 The Author(s)

All material available through ORBIT is protected by intellectual property law, including copyright law.

Any use made of the contents should comply with the relevant law.

[Deposit Guide](#)
Contact: [email](#)

RATIONALITY, PRAGMATICS, AND SOURCES

Rationality, Pragmatics, and Sources

Peter J. Collins

PhD Thesis

Psychological Sciences

Birkbeck, University of London

RATIONALITY, PRAGMATICS, AND SOURCES

This thesis is my own work. All collaborative aspects are appropriately indicated at the relevant points.

A handwritten signature in black ink, reading "Peter J. Hollan". The signature is written in a cursive, flowing style with a large initial 'P' and 'H'.

Abstract

This thesis contributes to the Great Rationality Debate in cognitive science. It introduces and explores a triangular scheme for understanding the relationship between rationality and two key abilities: pragmatics – roughly, inferring implicit intended utterance meanings – and learning from sources. The thesis argues that these three components – rationality, pragmatics, and sources – should be considered together: that each one informs the others. The thesis makes this case through literature review and theoretical work (principally, in Chapters 1 and 8) and through a series of empirical chapters focusing on different parts of the triangular scheme. Chapters 2 to 4 address the relationship between pragmatics and sources, focusing on how people change their beliefs when they read a conditional with a partially reliable source. The data bear on theories of the conditional and on the literature assessing people’s rationality with conditionals. Chapter 5 addresses the relationship between rationality and pragmatics, focusing on conditionals ‘in action’ in a framing effect known as goal framing. The data suggest a complex relationship between pragmatics and utilities, and support a new approach to goal framing. Chapter 6 addresses the relationship between rationality and sources, using normative Bayesian models to explore how people respond to simple claims from sources of different reliabilities. The data support a two-way relationship between claims and source information and, perhaps most strikingly, suggest that people readily treat sources as ‘anti-reliable’: as negatively correlated with the truth. Chapter 7 extends these experiments to test the theory that speakers can guard against reputational damage using hedging. The data do not support this theory, and raise questions about whether trust and vigilance against deception are prerequisites for pragmatics. Lastly, Chapter 8 synthesizes the results; argues for new ways of understanding the relationships between rationality,

RATIONALITY, PRAGMATICS, AND SOURCES

pragmatics, and sources; and relates the findings to emerging formal methods in psychology.

Table of Contents

Table of Contents	5
List of Figures	11
List of Tables.....	16
Acknowledgements	19
1 Theory Chapter.....	20
1.1 The Rationality Debate	22
1.1.1 Psychological Positions on Rationality.....	23
1.1.2 Detecting (Ir-)rationality in Experiments.....	34
1.2 Pragmatics and Rationality	37
1.2.1 Defining Pragmatics.....	37
1.2.2 Pragmatics and Theories of Rationality	45
1.2.3 Pragmatics and Experiments on Rationality	47
1.3 Sources	64
1.4 Prospectus	72
2 Conditionals and Testimony	74
2.1.1 Conditionals in the psychology of reasoning.....	76
2.1.2 Conditionals in Psycholinguistics	82
2.1.3 Conditionals and Pragmatics.....	82
2.1.4 Testimony.....	86
2.1.5 Towards an experimental method	87
2.2 Experiment 2.1: effect of assertion	90
2.2.1 Methods.....	90
2.2.2 Results & Discussion	93
2.3 Experiment 2.2: replicating the effect of Assertion	97

RATIONALITY, PRAGMATICS, AND SOURCES

2.3.1	Method	97
2.3.2	Results & Discussion	99
2.4	Experiment 2.3: effect of source expertise.....	104
2.4.1	Method	104
2.4.2	Results & Discussion	106
2.5	Experiment 2.4: replicating the effect of source expertise.....	109
2.5.1	Method	109
2.5.2	Results & Discussion	110
2.6	General Discussion.....	114
3	Conditionals, Testimony, and Interval Estimates	120
3.1	Experiment 3.1: Assertion and Interval Estimates.....	120
3.1.1	Method	120
3.1.2	Results and Discussion.....	123
3.2	Experiment 3.2: Expertise and Interval Estimates	130
3.2.1	Method	130
3.2.2	Results & Discussion	130
3.3	General Discussion.....	137
4	Independent Testimony, Priors, and a Model of Testimonial Conditionals	140
4.1.1	Multiple Testimony.....	140
4.1.2	Priors	142
4.2	Experiment 4.1: Independent testimony	144
4.2.1	Method.....	144
4.2.2	Results	145
4.2.3	Discussion	146

RATIONALITY, PRAGMATICS, AND SOURCES

4.3	Experiment 4.2: Priors, Assertion, and the probability of the antecedent.....	149
4.3.1	Methods.....	149
4.3.2	Results & Discussion.....	151
4.4	Experiment 4.3: Priors, Expertise, and the probability of the antecedent.....	154
4.4.1	Method.....	154
4.4.2	Results & Discussion.....	154
4.5	Experiment 4.4: Priors, Assertion, and the probability of the consequent.....	157
4.5.1	Methods.....	157
4.5.2	Results & Discussion.....	158
4.6	Experiment 4.5: Priors, Expertise, and the probability of the consequent.....	160
4.6.1	Methods.....	160
4.6.2	Results & Discussion.....	161
4.7	General Discussion.....	163
4.7.1	Modelling testimonial conditionals.....	166
4.7.2	Probabilities and Beyond.....	180
4.7.3	The meaning of the conditional.....	182
5	Framing and utility conditionals.....	188
5.1.1	Prospect Theory and the Asian Disease Paradigm.....	189
5.1.2	Towards goal framing.....	190
5.1.3	Goal framing: the empirical data.....	194

RATIONALITY, PRAGMATICS, AND SOURCES

5.1.4	Towards an experimental paradigm: utilities in argumentation and pragmatics	201
5.2	Experiment 5.1	206
5.2.1	Methods.....	208
5.2.2	Results	211
5.2.3	Discussion	216
5.3	Experiment 5.2	218
5.3.1	Methods.....	218
5.3.2	Results	219
5.3.3	Discussion	225
5.4	Experiment 5.3	230
5.4.1	Methods.....	231
5.4.2	Results	232
5.4.3	Discussion	236
5.5	General Discussion.....	239
6	Testimony and Source Reliability	244
6.1.1	Introducing Testimony	244
6.2	Experiment 6.1: Belief change.....	254
6.2.1	Method	254
6.2.2	Results & Discussion	258
6.3	Experiment 6.2: Replicating belief change.....	263
6.3.1	Methods.....	263
6.3.2	Results & Discussion	264
6.4	Experiment 6.3: Reliability change.....	266
6.4.1	Methods.....	266

RATIONALITY, PRAGMATICS, AND SOURCES

6.4.2	Results & Discussion	267
6.5	Experiment 6.4: Replicating reliability change.....	269
6.5.1	Methods.....	269
6.5.2	Results & Discussion	269
6.6	Experiment 6.5: Story version.....	272
6.6.1	Methods.....	274
6.6.2	Results & Discussion	275
6.7	General Discussion.....	278
7	Evidential Language.....	287
7.1	Experiment 7.1	290
7.1.1	Methods.....	290
7.1.2	Results	292
7.1.3	Discussion	294
7.2	Experiment 7.2	294
7.2.1	Methods.....	295
7.2.2	Results	296
7.2.3	Discussion	298
7.3	Experiment 7.3	298
7.3.1	Methods.....	298
7.3.2	Results	299
7.3.3	Discussion	301
7.4	Experiment 7.4: Shield Hedges.....	302
7.4.1	Methods.....	302
7.4.2	Results	303
7.4.3	Discussion	305

RATIONALITY, PRAGMATICS, AND SOURCES

7.5	General Discussion.....	305
7.5.1	Conclusions.....	311
8	General Discussion.....	313
8.1	The Experimental Data Recapitulated	313
8.1.1	Conditionals and Testimony	313
8.1.2	Utility Conditionals.....	315
8.1.3	Testimony and Source Reliability	317
8.1.4	Evidential Language.....	317
8.2	Redefining the relationships.....	318
8.2.1	Pragmatic intrusion into argumentation.....	321
8.2.2	Argument in pragmatics.....	327
8.2.3	Pragmatics as rational social action	338
8.3	Conclusions.....	343
	Bibliography.....	345

List of Figures

Figure 1.1. The Triangular Scheme.....21

Figure 1.2. Crupi's model34

Figure 2.1. Mean rating of P(Antecedent) by condition; error bars are standard error
.....94

Figure 2.2. Mean rating of P(Consequent) by condition; error bars are standard error
.....95

Figure 2.3. Mean rating of Conditional Probability by condition; error bars of
standard error96

Figure 2.4. Mean P(Antecedent) by condition; error bars are standard error99

Figure 2.5. Mean P(Consequent) by condition; error bars are standard error..... 101

Figure 2.6. Mean Conditional Probability by condition; error bars are standard error
..... 102

Figure 2.7. Mean P(Antecedent) by condition; error bars are standard error 106

Figure 2.8. Mean P(Consequent) by condition; error bars are standard error..... 107

Figure 2.9. Mean Conditional Probability by condition; error bars are standard error
..... 108

Figure 2.10. Mean P(Antecedent) by condition; error bars are standard error 110

Figure 2.11. Mean P(Consequent) by condition; error bars are standard error..... 112

Figure 2.12. Mean Conditional Probability by condition; error bars are standard error
..... 113

Figure 2.13. Simple Bayesian belief network for a conditional..... 117

Figure 3.1. Illustrative slider for exact estimate of $P=.5$ 121

Figure 3.2. Illustrative slider for uncertain estimate 122

RATIONALITY, PRAGMATICS, AND SOURCES

Figure 3.3. Mean point value of P(Antecedent) by condition; error bars are standard error	123
Figure 3.4. Mean slider range by condition; error bars are standard error.....	124
Figure 3.5. Mean point values of P(Consequent) by condition; error bars are standard error	125
Figure 3.6. Mean range of P(Consequent) by condition; error bars are standard error	126
Figure 3.7. Mean point values of Conditional Probability by condition; error bars are standard error	127
Figure 3.8. Mean range of Conditional Probability by condition; error bars are standard error	128
Figure 3.9. Mean point values of P(Antecedent) by condition; error bars are standard error	131
Figure 3.10. Mean slider ranges for P(Antecedent) by condition; error bars are standard error	132
Figure 3.11. Mean point values of P(Consequent) by condition; error bars are standard error	133
Figure 3.12. Mean slider ranges for P(Consequent) by condition; error bars are standard error	134
Figure 3.13. Mean point values for Conditional Probability by condition; error bars are standard error.....	135
Figure 3.14. Mean slider ranges for Conditional Probability by condition; error bars are standard error.....	136
Figure 4.1. Mean Conditional Probability by condition; error bars are standard error	145

RATIONALITY, PRAGMATICS, AND SOURCES

Figure 4.2. Effect of Assertion and Prior on P(Antecedent); error bars are standard error	151
Figure 4.3. Effect of Prior and Expertise on P(Antecedent); error bars are standard error	155
Figure 4.4. Effects of Prior and Assertion on P(Consequent); error bars are standard error	158
Figure 4.5. Effects of Prior and Expertise on P(Consequent); error bars are standard error	161
Figure 4.6. Bayesian belief network from Bovens and Hartmann (2003)	166
Figure 4.7. Baseline Bayesian belief network for experimental data	169
Figure 4.8. Comparison Bayesian belief network for experimental data.....	174
Figure 5.1. Argument convincingness by frame; error bars are standard error	211
Figure 5.2. Descriptive data by Utility and Frame. Circles are data points for positive frames; triangles are data points for negative frames. The solid line is the regression for positive frames; the dotted line for negative frames. Shading represents the standard error.	215
Figure 5.3. Descriptive data by Utility and Frame. Circles are data points for positive frames; triangles are data points for negative frames. The solid line is the regression for positive frames; the dotted line for negative frames. Shading represents the standard error.	220
Figure 5.4. Posterior distributions of mean utilities of taking medicine (left) and not taking medicine (right).....	222
Figure 5.5. Posterior distributions of mean utilities of undergoing surgery (left) and not undergoing surgery (right)	223

RATIONALITY, PRAGMATICS, AND SOURCES

Figure 5.6. Posterior distributions of means of utilities for risk of minor infection decreasing (left) and not decreasing (right)	224
Figure 5.7. Posterior distributions of mean utilities for risk of major illness decreasing (left) and not decreasing risk (right)	224
Figure 5.8. Mean utilities by condition; error bars are standard error	232
Figure 6.1. Mean belief change by reliability; error bars are standard error	261
Figure 6.2. Posterior distribution of effect size of belief change from reliable sources. ROPE from -.1, to .1; dotted lines are 95% HDI	261
Figure 6.3. Posterior distribution of effect size of belief change from unreliable sources. ROPE from -.1 to .1; black bar depicts 95% HDI.....	262
Figure 6.4. Mean belief change by reliability; error bars are standard error	264
Figure 6.5. Posterior distribution of effect size of belief change from reliable sources. ROPE from -.1 to .1; black bar depicts 95% HDI.....	265
Figure 6.6. Posterior distribution of effect size of belief change from unreliable sources. ROPE from -.1 to .1; black bar depicts 95% HDI.....	265
Figure 6.7. Mean change in reliability by expectedness; error bars are standard error	267
Figure 6.8. Posterior distribution of effect size of reliability change from expected claims. ROPE is -.1 to .1; black bar depicts 95% HDI	268
Figure 6.9. Posterior distribution of effect size of reliability change from unexpected claims. ROPE is -.1 to .1; black bar depicts 95% HDI	268
Figure 6.10. Mean reliability change by expectedness; error bars are standard error	270
Figure 6.11. Posterior distribution of effect size of reliability change from expected claims. ROPE from -.1 to .1; black bar depicts 95% HDI	270

RATIONALITY, PRAGMATICS, AND SOURCES

Figure 6.12. Posterior distribution of effect size of reliability change from unexpected claims. ROPE from -.1 to .1; black bar depicts 95% HDI	271
Figure 6.13. Mean belief by expectedness; error bars are standard error	276
Figure 6.14. Posterior distribution of effect size of difference between expected and null conditions. ROPE from -.1 to .1; black bar depicts 95% HDI.....	277
Figure 6.15. Posterior distribution of effect size of difference between unexpected and null claims. ROPE from -.1 to .1; black bar is 95% HDI.....	277
Figure 6.16. Posterior distribution for effect size of difference between expected and unexpected conditions. ROPE from -.1 to .1; black bar depicts 95% HDI.....	278
Figure 7.1. Mean change scores by condition; error bars are standard error	293
Figure 7.2. Mean change scores by condition; error bars are standard error	297
Figure 7.3. Mean change score by condition; error bars are standard error.....	299
Figure 7.4. Mean claim strength; error bars are standard error.....	301
Figure 7.5. Mean change scores by condition; error bars are standard error	303
Figure 7.6. Mean claim strength (manipulation check); error bars are standard error	304

List of Tables

Table 2.1. Truth table of material conditional.....	77
Table 2.2. Defective truth table.....	78
Table 2.3. Fixed effects of Assertion on P(Antecedent)	95
Table 2.4. Fixed effects of Assertion on P(Consequent)	96
Table 2.5. Fixed effect of Assertion on Conditional Probability	97
Table 2.6. Fixed effect of Assertion on P(Antecedent).....	100
Table 2.7. Fixed effect of Assertion on P(Consequent).....	101
Table 2.8. Fixed effects of Assertion on Conditional Probability.....	103
Table 2.9. Fixed effects of Expertise on P(Antecedent)	107
Table 2.10. Fixed effects of Expertise on P(Consequent).....	107
Table 2.11. Fixed effects of Expertise on Conditional Probability.....	108
Table 2.12. Fixed effects of Expertise on P(Antecedent)	111
Table 2.13. Fixed effects of Expertise on P(Consequent).....	112
Table 2.14. Fixed effects of Expertise on Conditional Probability.....	113
Table 3.1. Fixed effects of Assertion on point values of P(Antecedent)	124
Table 3.2. Fixed effects of Assertion on slider ranges of P(Antecedent)	125
Table 3.3. Fixed effects of Assertion on point values of P(Consequent).....	126
Table 3.4. Fixed effects of Assertion on slider ranges for P(Consequent)	127
Table 3.5. Fixed effects of Assertion on Conditional Probability.....	128
Table 3.6. Fixed effects of Assertion on ranges for Conditional Probability	129
Table 3.7. Fixed effects of Expertise on point values of P(Antecedent).....	131
Table 3.8. Fixed effects of Expertise on slider ranges for P(Antecedent)	132
Table 3.9. Fixed effects of Expertise on point values for P(Consequent).....	133

RATIONALITY, PRAGMATICS, AND SOURCES

Table 3.10. Fixed effects of Expertise on slider ranges for P(Consequent); error bars are standard error.....	134
Table 3.11. Fixed effects of Expertise on point values for Conditional Probability.....	135
Table 3.12. Fixed effects of Expertise on slider ranges for Conditional Probability....	136
Table 4.1. Fixed effects of Assertion on Conditional Probability.....	146
Table 4.2. Fixed effects and confidence intervals of single assertion for the point-estimate studies	147
Table 4.3. Fixed effects of Prior, Assertion, and Interaction	152
Table 4.4. Fixed effects of Assertion at low prior.....	152
Table 4.5. Fixed effects of Assertion at high prior	153
Table 4.6. Fixed effects of Prior, Expertise, and Interaction on P(Antecedent)	155
Table 4.7. Fixed effects of Expertise on P(Antecedent) at low prior.....	156
Table 4.8. Fixed effects of Expertise on P(Antecedent) at high prior.....	156
Table 4.9. Fixed effects of Prior, Assertion, and Interaction on P(Consequent)	159
Table 4.10. Fixed effects of Assertion on P(Consequent) at low prior.....	159
Table 4.11. Fixed effects of Assertion on P(Consequent) at high prior.....	160
Table 4.12. Fixed effects of Prior, Expertise, and Interaction on P(Consequent).....	162
Table 4.13. Fixed effects of Expertise on P(Consequent) at low prior	162
Table 4.14. Fixed effects of Expertise on P(Consequent) at high prior.....	163
Table 4.15. Conditional probability table for Bovens Hartmann network.....	167
Table 4.16. Conditional probability table for node B	170
Table 4.17. Conditional probability table for node Rep X.....	170
Table 4.18. Conditional probability table for Rep X	174
Table 5.1. Utilities for goal frames.	206
Table 5.2. Fixed effects of Frame on Convincingness.....	213

RATIONALITY, PRAGMATICS, AND SOURCES

Table 5.3. Counts for antecedent utilities.....	213
Table 5.4. Counts for consequent utilities; 'Minor Protection' is protection against a minor infection, 'Major Protection' against a life-threatening illness	214
Table 5.5. Fixed effects of Utility, Frame, and Interaction on Convincingness	216
Table 5.6. Fixed effects of the full additive model	216
Table 5.7. Fixed effects of Frame, Utility, and Interaction.....	221
Table 5.8. Fixed effects for full additive model.....	221
Table 5.9. Parameters for model with interaction	234
Table 8.1. Dialogue types according to Walton (2008, p. 8)	322

Acknowledgements

Some words of thanks are due. Firstly, thanks are due to Professor Ulrike Hahn for her tremendous supervision over the past three years – for her remarkable support and encouragement. I would also like to thank Professor Mike Oaksford for helpful discussion of the results of data from Chapters 2 to 4. Since these chapters resulted from collaborative work, I would like to thank these collaborators: Dr Karolina Krzyżanowska and Professors Stephan Hartmann and Gregory Wheeler. Dr Jean-François Bonnefon provided invaluable suggestions during the planning of the studies in Chapter 5, and gave stimulating feedback on the results in that chapter. Chapter 6 resulted from collaborative work with Dr Ylva von Gerber and Professor Erik Olsson; they, also, deserve thanks. I have benefitted from fruitful discussions on both conditionals and pragmatics with Drs Karolina Krzyżanowska and Niels Skovgaard -Olsen, and from helpful advice on mixed-effects modelling from Dr Henrik Singmann and Danielle Pessach. Ruben Zamora gave invaluable help and advice with the web experiments, and Joe Miele, of MTurk Data, posted the web experiments on Mechanical Turk with great efficiency. I also gratefully acknowledge financial support from the Bloomsbury Doctoral Training Centre and experimental funding from the Alexander von Humboldt Foundation and the Swedish Research Council. Last – but by no means least – I thank Matt Evans for constant love and support, and for well-timed advice to ‘Have a word with [myself]!’

1 Theory Chapter

The question ‘How rational are we?’ has stimulated a rich debate in cognitive science. Cognitive scientists argue over the norms that define optimal behaviour; over whether we form and revise our beliefs optimally; over whether we act optimally to achieve our goals; and over whether we are educable (Stanovich, 2012). This thesis will contend that rationality is inextricably linked with two abilities: pragmatics - our ability to interpret utterances, including their implicit speaker-intended meanings; and our ability to respond to information that we receive from sources, where sources are understood as other people or institutions that are intentionally giving information. The novelty is not in suggesting a role for pragmatics and sources: the existing literature has already suggested as much. The novelty is, rather, to stress the deep relationship among the three components: rationality, pragmatics, and sources. The thesis will contend that much can be learnt by treating these components together: indeed, that we lose valuable insights if we consider the components separately.

Throughout, this thesis will refer to a triangular scheme to guide discussion. Figure 1.1 shows this scheme.

RATIONALITY, PRAGMATICS, AND SOURCES

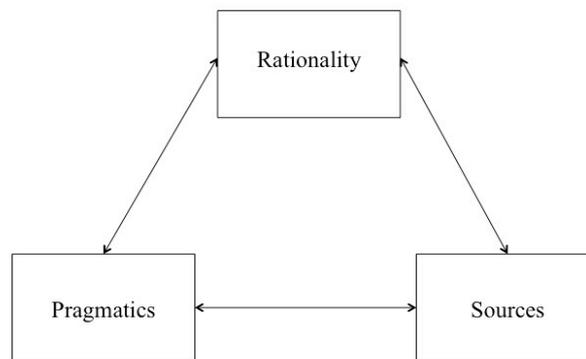


Figure 1.1. The Triangular Scheme

The figure represents the three components and the two-way relationships among them. The arrows can be taken to represent both conceptual connections and strategies: for example, because rationality and pragmatics are fundamentally linked, the study of rationality is illuminated by considering pragmatics and, likewise, the study of pragmatics is illuminated by considering rationality.

This thesis will present both conceptual arguments and empirical data to make the case for this triangular scheme. The thesis is structured around four sets of experiments. The experiments offer novel data which illustrate a different aspect of the triangular scheme. The experiments also contribute to four independent debates in cognitive science: on learning conditionals from testimony; on utility conditionals; on models of testimony; and on evidential language. The experiments are tied together by the triangular scheme.

This introductory chapter will introduce the rationality debate in cognitive science and how the two other components, pragmatics and sources, bear on this debate. The chapter will inevitably be slanted towards topics on which there is extensive data: principally, rationality in general, and the relationship between

rationality and pragmatics. This slant reflects the state of the literature, not the importance of the components.

1.1 The Rationality Debate

Rationality arises as a topic across psychology. Given this wide potential range, some focus is needed. Consider, for instance, ideal observer models of perception, where formal models of optimal performance are specified and compared with human behaviour (for discussion, see Hahn, 2014). Similar rational models exist for numerous other domains, such as categorization (Anderson, 1991b) and language acquisition (Perfors, Tenenbaum, & Regier, 2011). Rationality also bears on the definition and treatment of mental disorders (Graham, 2014). But its most familiar territory is a cluster of related sub-disciplines: the psychology of judgment and decision making (see, e.g., Newell, Lagnado, & Shanks, 2015); the psychology of reasoning (see, e.g., Manktelow, 2012); and the psychology of argumentation (see, e.g., Hahn & Oaksford, 2007). It is these disciplines which will be the focus of this thesis.

When assessing rationality – whatever normative standards we choose – the crucial evidence is not momentary successes and failures but systematic rationality or irrationality. In other words, the key question is our underlying competence. Establishing our competence requires abstracting away from performance imitations, such as memory capacity¹, tiredness, and so on (Stein, 1996). Contemporary psychology seems to assume what Stein (1996, p. 12) calls the Irrationality Thesis:

¹ Note the tension with accounts which factor in computational limitations (see the section on Bayesianism and Rational Analysis). One way to try to reconcile these approaches is to distinguish between hard computational limits, such as absolute limits on working memory, and contextual problems, such as those caused by tasks with heavy working-memory load.

our competence systematically diverges from normative principles. The meat of the debate is the extent of this divergence.

1.1.1 Psychological Positions on Rationality.

We can classify the most prominent theories by the normative frameworks that they endorse² and their degree of optimism about rationality (V. Crupi, personal communication, September 5, 2016).

The Standard Picture.

Probably the dominant normative framework is the ‘Standard Picture’ (Stein, 1996): the correct norms are the ‘rules of logic, probability and so forth’ (Stein, 1996, p. 4). The Standard Picture has shaped much of the literature surveyed below. Implicit in the framework is the understanding that ‘logic’ and ‘probability’ refer to classical accounts: classical propositional and predicate logic, and probability governed by the Kolmogorov axioms (Kolmogorov, 2013; for discussion of classical probability and an alternative, see Busemeyer & Bruza, 2012). Implicit, too, is the inclusion of prescriptions from decision theory: in particular, the maximization of expected utility. For Stein (1996) refers extensively to a decision-making literature which typically proceeds by assessing adherence to the axioms of choice, full adherence to which would imply maximization of expected utility (Stanovich, 2016).

Bayesianism. There are many and varied Bayesian models within cognitive science: too many to address individually. But many of these models are drawn

² This section is inspired by Vincenzo Crupi’s lecture series, ‘Norms versus Reasoning’, delivered from September 5th to September 9th, 2016, at the International Rationality Summer Institute, in Aurich, during which Crupi identified the leading relevant theories and classified them according to the norms endorsed and the degree of optimism about rationality. I largely preserve Crupi’s choice of theories, but add my own analysis of

RATIONALITY, PRAGMATICS, AND SOURCES

together in the account of Oaksford and Chater (e.g. 2007, 2009), which they call ‘Bayesian Rationality’. Bayesian Rationality assumes the Standard Picture of rationality. Norms are stated in terms of probabilities, understood as subjective degrees of belief (Oaksford & Chater, 2007; on different interpretations of probability, see, e.g., Hacking, 2001). This preference for probability does not imply a conflict with logic-based approaches, since probability and logic are tightly connected (see, e.g., Busemeyer & Bruza, 2012). Indeed, the Bayesian probabilistic account can be seen as a generalization of logic to handle real-world uncertain reasoning (Oaksford & Chater, 2009b). With this broader account, Bayesian Rationality allows a unified treatment of judgment, decision making, inductive and deductive reasoning, and argumentation (Oaksford & Chater, 2009a,b).

Bayesian Rationality draws heavily on the program of rational analysis (see, e.g., Anderson, 1990, 1991a). As presented in Oaksford and Chater (2007, p. 32), rational analysis comprises the following steps:

1. Specify precisely the goals of the cognitive system.
2. Develop a formal model of the environment to which the system is adapted.
3. Make minimal assumptions about computational limitations.
4. Derive the optimal behaviour function given 1-3 above. (This requires formal analysis using rational norms, such as probability theory and decision theory).
5. Examine the empirical evidence to see whether the predictions of the behaviour function are confirmed.

the theories and, consequently, my own assessment of their degree of optimism about

6. Repeat, iteratively refining the theory.

The framework has a crucial normative component, but judges norms only after considering goals, environment and computational limitations.

Bayesian Rationality makes a largely optimistic assessment of rationality. In various domains, after rational analysis, people's competence is said to approximate rational principles (e.g. Oaksford & Chater, 2007, 2009). For instance, there have been striking reassments of apparent irrationality in deductive and syllogistic reasoning (Oaksford & Chater, 1994; Oaksford, Chater, & Larkin, 2000) and informal argumentation (Hahn & Oaksford, 2007). Nevertheless, irrationality is not simply explained away. Indeed, it is explicitly acknowledged that 'people's ability to deal with verbally (or, in the case of probabilities, numerically) specified reasoning problems...is extremely limited and subject to error' (Oaksford & Chater, 2007, p. 13) and that their probabilistic reasoning is primarily qualitative rather than quantitative (Oaksford & Chater, 2009). Consequently, there is no need to dispute classic findings demonstrating irrationality, such as conservatism in belief revision (e.g. Phillips & Edwards, 1966), in precisely the domains where one might expect a rational Bayesian agent to excel. Rather such findings arguably show that, in Bayesian terms, people respond to evidence in qualitatively appropriate, but quantitatively inappropriate, ways (Hahn & Harris, 2014). Thus, in Stein's (1996) terms, even optimistic Bayesianism is a variant of the Irrationality Thesis.

Mental Logic and Mental Models. These two approaches are both inspired by classical logic, and largely focus on the psychology of reasoning. Mental logic takes a syntactic (i.e. rule-based) approach to reasoning, developed from natural deduction,

rationality. This assessment should not, therefore, be taken to reflect Crupi's own views.

RATIONALITY, PRAGMATICS, AND SOURCES

a method of formal-logical proof (e.g. Braine & O'Brien, 1991; Braine & O'Brien, 1998; Rips, 1994; on natural deduction, see, e.g., Arthur, 2011). Although mental logic assumes formal logic as a standard, there is no direct correspondence between formal and mental logical rules. Unlike formal logic, for instance, Braine and O'Brien's (1991) mental logic does not license *modus tollens* (if p then q; not q; therefore not p). Such inferences still occur, but through strategic thinking by able reasoners (Manktelow, 2012). Other differences occur because of processing constraints: reasoning can break down because of schema complexity (Braine & O'Brien, 1991) or the number of rules called on (Rips, 1994). Similarly, contextual information can change the accessibility of rules. For example, when the context is permission being granted, denial of the antecedent becomes more accessible ('if p, then q; therefore if not p, then not q'; see Braine & O'Brien, 1991).

Mental Models Theory takes a more semantic (or content-based) approach. The theory resists summary, because it continues to undergo substantial changes³. But a stable point is that Mental Models Theory assumes classical logic as a normative standard: 'to be rational is to be able to make deductions – to draw valid conclusions from premises. A valid conclusion is one that is true in any case in which the premises are true' (Johnson-Laird, Khemlani, & Goodwin (2015, p. 201). Non-normative patterns routinely arise, the reasons for which recall Mental Logic above. Firstly, there are differences in representation. Mental models are concrete representations of particular situations, and are economical. People initially represent only true information, leaving them prone to certain fallacies, known as illusory

inferences (Johnson-Laird, Khemlani, & Goodwin, 2015). People may flesh out models fully, but this ability depends on factors such as age, intelligence, and working-memory load, the latter being a function of the number of models invoked (Manktelow, 2012). Reasoning is also modulated by content, context and knowledge (Johnson-Laird & Byrne, 2002; Johnson-Laird, Khemlani, & Goodwin, 2015). People can add information to models, such as causal relations between the antecedents and consequents of conditionals (Johnson-Laird & Byrne, 2002), and they can use their background knowledge to eliminate models. For instance, on representing ‘Pat visited Milan or she visited Italy’, people will eliminate the possibility that Pat visited Milan and not Italy (Johnson-Laird, Khemlani & Goodwin, 2015). Actual concrete reasoning can be far richer than the prescriptions of formal logic (Johnson-Laird & Byrne, 2002).

Dual-process Theories. It has been a dominant trend in the psychology of judgment, decision making and reasoning to posit two types of processing, referred to as Type 1 (System 1) and Type 2 (System 2). This approach assumes the Standard Picture, and argues for considerable mismatch between norms and actual behaviour. The distinction between Type 1 and Type 2 plays a major role in accounting for this mismatch. The distinction can be fleshed out in numerous ways (for discussion, see Evans & Stanovich, 2013, and especially Stanovich, 2012). There are, however, some emerging commonalities (Stanovich, 2012). Type 1 thinking corresponds to The Autonomous Set of Systems (Stanovich, 2005): systems which operate autonomously, have their own triggering stimuli, and are not under higher level

³ Compare, for instance, the treatment of logical connectives in Johnson-Laird & Byrne (1991), Johnson-Laird & Byrne (2002) and Johnson-Laird, Khemlani, and Goodwin (2015)(see, also, Baratgin et al., 2015, and Evans & Over, 2004, for critical discussion).

RATIONALITY, PRAGMATICS, AND SOURCES

cognitive control (Stanovich, 2012). Type 1 processing also tends to be fast and computationally cheap, and can operate in parallel (Stanovich, 2012). In contrast, Type 2 processing is not autonomous, and tends to be slow, computationally expensive and serial (Stanovich, 2012).

For dual-process theories, individual differences are key to assessing rationality. Type 1 processing is largely uniform, excepting pathological variation, such as autism, agnosia and alexia (Stanovich, 2012). Type 1 processing provides a baseline of (ir-)rationality: behaviours that work well in some contexts but give rise to systematic biases in others (for discussion of these contexts, see Kahneman & Klein, 2009). In contrast, Type 2 processing is markedly varied. People can show profound differences in fluid intelligence, that is, in their ability to solve novel problems (Stanovich, 2012). They also show profound differences in reflective thinking: for instance, in their tendency to gather information and think extensively before reaching a conclusion; judge degrees of evidence and avoid absolute judgments; and predict consequences and consider their likely costs and benefits. These thinking dispositions – and fluid intelligence – are given much credit for determining an individual’s rationality, since both contribute independently to performance on standard decision and judgment problems (Stanovich, 2012). But crucial, too, is the balance between Type 1 and Type 2 processing⁴. On this, the literature differs, but leading accounts emphasize that Type 2 processing is slow, effortful, and lazy (Evans & Stanovich, 2013; Kahneman, 2011; Stanovich, 2012; Stanovich & West, 1999).

⁴ Note, though, that some accounts, such as Thompson (2010), caution that normative responses can result from Type 1 or Type 2 processing.

Beyond the Standard Picture.

By no means all psychological theories cleave to the Standard Picture of rationality. There are ample opportunities to diverge from classical logic, probability and decision theory.

Stenning and van Lambalgen. A leading theory gives logic, or rather logics, a central role in the psychology of reasoning (Stenning & Lambalgen, 2012), but these are not the classical logics of the Standard Picture. Logical reasoning is taken to be ‘simultaneously formal and relative to a domain’ (Stenning & van Lambalgen, 2012, p. 19). Logical reasoners must first reason to an interpretation of the task in hand, and then reason from that interpretation. Reasoning to an interpretation amounts to parameter setting, the parameters being a choice of formal language, its semantics, and its definition of validity (Stenning & van Lambalgen, 2012, p. 25). For example, the interpretation would be classical logic if the chosen formal logic is recursive and truth-functional, with a bivalent semantics, and with validity understood as true on all models (Stenning & van Lambalgen, 2012). But bivalence could be dropped, resulting in a trivalent or even fuzzy logic (Stenning & van Lambalgen, 2012). Reasoning from an interpretation amounts to deriving valid conclusions as defined by the selected logic.

For Stenning and van Lambalgen (2012, p. 7), ‘reasoning is everywhere’ and interpretation is key. Consequently, many cases of apparent irrationality can, in principle, be explained away by invoking a different logic. The use of these logics can be empirically supported by evidence from, for example, Socratic questioning (for discussion, see Stenning and van Lambalgen, 2012). It is less clear that these different logics can be normatively supported. We might suggest, for instance, that reasoning can be modelled in many contexts as deductive reasoning with exceptions:

RATIONALITY, PRAGMATICS, AND SOURCES

that is, reasoning of the form ‘If P and No Exceptions, then Q’. But if such a logic is to be normatively binding, it will need a comprehensive normative theory of what counts as an exception. We do not have a promising theory (Hahn & Hornikx, 2016).

Stenning and van Lambalgen (2012) are comparatively optimistic about human reasoning, but acknowledge the following failings. Firstly, we may fail to reason consistently from our own interpretation. Secondly, our primary skill may be in credulous reasoning – that is, in identifying a single interpretation that would make a speaker’s utterances true, as in discourse comprehension. Although we can be good at more sceptical reasoning – that is, in identifying conclusions which are true on all interpretations – we may require supportive contexts, which are absent in typical experiments on reasoning (for discussion, see Stenning & van Lambalgen, 2012, p. 357). Thirdly, we may be unaware of the interpretive process – of our ability to manipulate interpretations – and may consider too few interpretations.

Fast and Frugal Heuristics. This approach, pioneered by Gerd Gigerenzer and colleagues (see, e.g., Gigerenzer, 2000, 2010, 2015; Gigerenzer & Selten, 2002; Gigerenzer, Todd, & Group, 1999), focuses on judgement and decision making and, like Stenning and van Lambalgen (2012), rejects a role for content-free norms. The research program aims to identify people’s ‘adaptive toolbox’, their stock of ‘heuristics’ or simple solutions to a problem which require little information (Gigerenzer, 2010). While standard economic models assume that, for a given decision, people have a utility function which they attempt to optimize, heuristic models assume that people satisfice, that is, find a solution that is just good enough (see, e.g., Gigerenzer, 2010; Katsikopoulos, 2014). Heuristic models are assessed for their ecological rationality: the researcher asks when – in what environments – the

RATIONALITY, PRAGMATICS, AND SOURCES

heuristics are successful. In other words, the implied standard is simply what works in context.

The program is largely optimistic about rationality because of the apparent ecological rationality of heuristics, but it is a qualified optimism. An especially successful heuristic is the 1/N heuristic: the common investing strategy to divide an investment equally into a certain number of assets. This heuristic was tested against 14 optimization models in a simulation study (DeMiguel, Garlappi, & Uppal, 2009). No optimization model consistently outperformed the heuristic. To outperform it, the optimization models needed 3,000 months' data for a portfolio with 25 assets, and 6,000 months' data for a portfolio with 50 assets (DeMiguel, Garlappi & Uppal, 2009). Nevertheless, there are limits to rationality. At least under conditions of uncertainty, satisficing requires less information, does not need an optimal solution to exist, and is computationally cheaper (Gigerenzer, 2010). However, there may be circumstances (principally, decision making under risk⁵) where more classical strategies, such as optimization, are appropriate (Gigerenzer, 2010). Here, the program would have to accept that people may fail to optimize: that they fall short of normative behaviour. Even in decision making under uncertainty, though, the program implicitly acknowledges a shortfall. The research program has an applied dimension - the design of novel heuristics for helping people to solve problems better (Gigerenzer, 2010) – which presupposes that people are not perfectly rational.

⁵ In his lecture, 'The Ecological Rationality of Heuristics', at the International Rationality Summer Institute, Gigerenzer (personal communication, September 6, 2016) recently stressed that the conclusions of the 'Fast and Frugal Heuristics' program apply principally to decision making under uncertainty, and not to decision making under risk. Gigerenzer argued, however, that decision making under uncertainty is the larger, more significant domain.

No Norms.

As we have seen, there are diverse positions, in psychology, on how to define rationality and how rational people are. This diversity has led some to question the usefulness of the rationality debate. Elqayam and Evans (2011) argue for a reorientation of cognitive science away from norms and towards pure description of people's behaviour. Elqayam and Evans complain of an arbitration problem: for (almost) any single experimental test of rationality, there will be competing norms available; and the choice between these norms is both not obvious and not obviously part of psychology. Take, for instance, the Wason selection task, in which people are asked to test a rule such as 'If there is an A on one side of a card, then there is a 2 on the other side' (Wason, 1960). Should we apply the norms of propositional logic or of Bayesian information selection (see, e.g., Oaksford & Chater, 1994)? This question does not seem intrinsically psychological. Elqayam and Evans argue, further, that a focus on norms has led to damaging biases in research on thinking, such as the historical focus on logic and deduction in the psychology of reasoning. Norms, they suggest, may only be of interest if we aim, not just to study, but to improve thinking.

This position has, unsurprisingly, proved controversial, and has been met with persuasive responses. A straightforward response is that norms are debated in precisely those areas which we most likely want to improve: our judgments, decision making, reasoning and so on. It is no coincidence that research in these areas has heavily influenced 'nudge' programs internationally (for discussion, see Bonnefon, 2013; Gigerenzer, 2015; Thaler & Sunstein, 2009). More substantively, there is a robust argument that norms are, in fact, stimulating and productive. Douven (2011), for instance, gives the example of belief revision. If our goal is long-term accuracy,

RATIONALITY, PRAGMATICS, AND SOURCES

then we should update our beliefs in a Bayesian fashion. Do we do so? And if not, why not? Here, considering norms suggests a distinctive research program. Similarly Hahn (2011, 2014) argues that norms can guide our interpretation of behaviour. A given behaviour is compatible with many possible models. We can narrow the search space for these models if we make the reasonable assumption that ‘the cognitive system is trying to do something sensible’ and consider optimal models and approximations to them (Hahn, 2014, p. 10). A useful analogy, here, is with ideal-observer analysis in the study of perception, where an optimal model is specified, compared with human behaviour, and then refined iteratively by building in capacity limitations (Hahn, 2014).

This thesis will take the position that considering norms and rationality can, indeed, be richly productive, especially when using the framework in the following section. The thesis will also adopt Crupi’s (personal communication, September 5, 2016) response to the arbitration problem. We can grant that the study of rationality involves frequent cases of clashing norms. We can grant, too, that psychology is not the natural place for deciding between these norms, this being more naturally the business of philosophy (Crupi & Girotto, 2014). But a solution to the problem is implicit in the way this section has been framed: as the rationality debate within cognitive science. Cognitive science is inherently interdisciplinary, and includes both psychology and philosophy. Through interdisciplinary work and a division of labour, empirical research on rationality can proceed alongside conceptual work on the arbitration problem. The following section will outline Crupi’s model for how such research can proceed. Crupi’s model also prepares the ground for the introduction of pragmatics.

1.1.2 Detecting (Ir-)rationality in Experiments

In his lecture series ‘Norms versus Reasoning’, Vincenzo Crupi sketched a framework for assessing rationality in experimental tasks (see, also, Crupi & Girotto, 2014). The framework links norms and behaviour in a way that will prove invaluable for this thesis. Consider Figure 1.2 below.

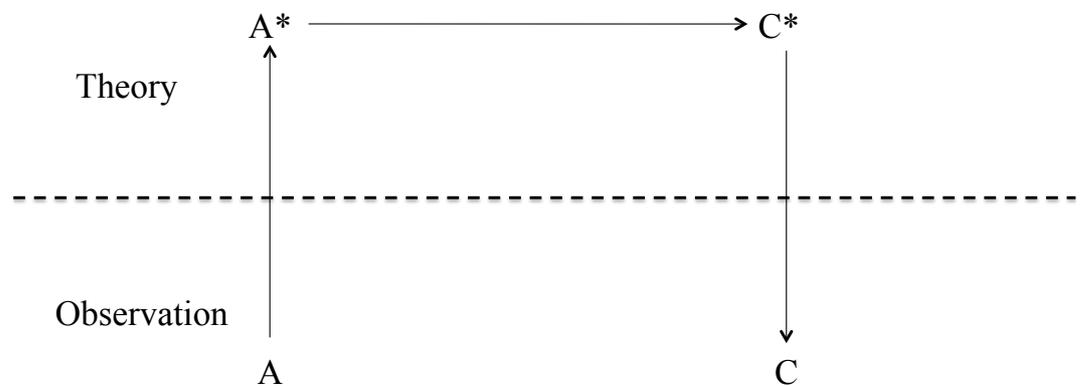


Figure 1.2. Crupi's model

The framework distinguishes between an observational and a theoretical level. At the observational level is the actual experiment. Participants are presented with a set of premises - the experimental instructions and materials - which they interpret in some way. These interpreted premises correspond to A in the diagram above. From these premises, participants derive conclusions and make their response using the options supplied in the experiment (e.g. a rating on a Likert-style scale, the selection of a one option in a forced choice, and so on). This response corresponds to C in the diagram above. The descriptive task is to develop a model to account for the transition from

RATIONALITY, PRAGMATICS, AND SOURCES

premises, A, to response, C. The theoretical level is where rationality is determined. Here there are premises, A*, and conclusions, C*. The normative task is to assess whether C* can correctly be derived from A*. To assess this, we need a normative system, generally specified in a formal language such as classical logic or the probability calculus. Consequently A* and C* are specified in the chosen formal language. In many cases, there will, as Elqayam and Evans (2011) observe, be a choice of norms; there will, in other words, be an arbitration problem. This choice will generally require philosophical, rather than psychological, argumentation (Crupi & Girotto, 2014).

In the Crupi framework, to adjudicate on rational behaviour in an experiment, we must consider the following questions (for discussion, see Crupi & Girotto, 2014). Has the appropriate normative framework been selected (the arbitration problem)? Assuming these norms, have they been correctly applied? Is there (mis-)match between participants' and experimenters' understanding of the premises (task instructions and materials)? Is there (mis-)match between participants' and experimenters' understanding of the conclusions (responses)? As Crupi and Girotto (2014) argue, these questions have generated productive debate in the psychological literature. We saw above that the arbitration problem has arguably stimulated considerable research and aided in the selection of correct models. Considering norm application has stimulated research on pseudo-diagnostics (e.g. Crupi, Tentori, & Lombardi, 2009; Tweney, Doherty, & Kleiter, 2010). Considering premise (mis-)match lies behind Oaksford and Chater's (1994) analysis of the Wason selection task: they suggest that participants interpret the conditional rule as referring to the conditional probability, not as a material conditional (for discussion, see Crupi & Girotto, 2014). Lastly, considering conclusion (mis-)match lies behind the extensive

RATIONALITY, PRAGMATICS, AND SOURCES

debate on the conjunction fallacy and how people interpret the apparently simple sentences ‘Linda is a bank-teller’ and ‘Linda is a bank-teller and a feminist’ (e.g. Hertwig, Benz, & Krauss, 2008; Hertwig & Gigerenzer, 1999; Jarvstad & Hahn, 2011; Tentori & Crupi, 2012; Tentori, Crupi, & Russo, 2013; Tversky & Kahneman, 1983).

The Crupi framework highlights the importance of interpretation.

Researchers do not generally present participants with premises and conclusions, stated in some formalism, and assess which conclusions participants endorse. Instead, researchers typically present experiments in natural language, and even tasks that are primarily non-linguistic need instructions in natural language (for discussion, see Lee, 2006). Interpretation features in the diagram above in the linking of A and A*, and C and C*, with arrows. These arrows reflect an evidential requirement for studies on rationality: the interpretations of participant and experimenter should match. Importantly, interpretation is neutral with respect to the chosen norms. That is, whichever norms a researcher might select, s/he must consider whether his/her interpretation of the task, premises and conclusions matches that of participants.

The Crupi framework raises the question of what interpretation means. In a broad sense, to interpret the language in an experiment, participants must apply much of their linguistic competence. At a minimum, they must map the written or spoken words onto morphology, word meanings, and syntax. Having such competence is a basic prerequisite for performing an experimental task, but this dependence is somewhat trivial. More interesting is the dependence on understanding the meaning of the utterances that comprise the experimental materials. At this point we introduce the second component in the triangular scheme: pragmatics.

1.2 Pragmatics and Rationality⁶

1.2.1 Defining Pragmatics

What do utterances mean? Meaning is typically separated into two components: semantics and pragmatics. The latter is especially important to rationality. Defining the two terms is not a simple matter (for discussion, see, Huang, 2007; Levinson, 1983), but a single watertight definition is not crucial for this thesis. It will suffice to adopt a working definition, constructed from the following common distinctions.

Sentence meaning vs. utterance meaning: semantics concerns the meanings of words and how these meanings compose to form the meanings of sentence, whereas pragmatics concerns the meaning of utterances.

When a sentence is issued in an actual context, it becomes an utterance (Levinson, 1983, p. 18). Trivially, the sentence ‘It will rain today’ acquires a different meaning when uttered on different days. But what aspects of the context are particularly important? Utterances are actions at a given time and place and they are, crucially, intentional actions by a speaker, which gives rise to the second distinction.

Sentence meaning vs. speaker meaning: semantics is sentence meaning; pragmatics is speaker meaning.

Although the sentence ‘This thesis is fascinating’ expresses a high degree of interest in a particular thesis, it can be intended ironically and, cued by appropriate intonation, the opposite meaning will be understood (example based on Levinson, 1983, p. 17). Mention of irony brings us to a third distinction (for discussion, see, e.g., (Korta & Perry, 2015) .

⁶ The text on the definition of pragmatics is adapted from Collins and Hahn (2016).

RATIONALITY, PRAGMATICS, AND SOURCES

Literal vs. non-literal meaning: semantics is literal meaning; pragmatics is non-literal meaning.

This distinction has intuitive appeal, although a formal definition of literalness is far from straightforward (Recanati, 2004). By ‘non-literal’ is meant not just figurative language, such as metaphor and irony, but any meaning which goes beyond the content of what is said. A final distinction draws on differences in the abilities underlying semantics and pragmatics (for discussion, see B. Clark, 2013; Sperber & Wilson, 1995).

Decoding vs. inference: Semantics is decoding meaning; pragmatics is inferring it.

This distinction takes the hearer’s perspective; for the speaker’s, replace ‘decoding’ with ‘encoding’ and ‘inferring’ with ‘implying’. These four distinctions are complementary, and can straightforwardly be combined into the following working definition:

Semantics: decoding (encoding) the meaning of words, sentences, and utterances⁷.

Pragmatics: inferring (implying) the implicit, speaker-intended meaning of an utterance.

The definition of pragmatics excludes non-literalness because of the complexities of defining the term robustly (Recanati, 2004). In many cases, pragmatic meanings will be intuitively non-literal.

This working definition makes pragmatics heavily reliant on theory of mind, a point which already indicates the close relationship between pragmatics and

information about message sources (speakers). It has been a widespread view since Grice (1957) that successful communication depends on intention recognition. The intention in question is the speaker's intention towards an addressee: for instance, in the case of assertion, the speaker intends the addressee to believe the asserted proposition. This intention is overt: it is supposed to be recognized by the hearer. How to formalize such communicative intentions is widely debated (see, e.g., Breheny, 2006; Clark, 1996; Grice, 1957; Schiffer, 1972; Sperber & Wilson, 1995; Strawson, 1964). By consensus, though, overtness is paramount (Korta & Perry, 2015), and communication is taken to depend on at least first-order, perhaps also second-order, theory of mind (Breheny, 2006; Sperber & Wilson, 1995).

Although intention recognition helps to pin down communication, it does not explain how people actually go about deriving pragmatic meanings. Grice (1975) offered a plausible account, which gives insight into how experimental materials might be understood. The Gricean account comprises an overarching Cooperative Principle and attendant Maxims of Conversation. The following principle and maxims can be construed as descriptions of both how a hearer expects the speaker to behave and of a speaker's conversational aims.

Cooperative Principle: "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged." (Grice, 1975, p. 45)

Maxims:

Quantity:

⁷ Including 'utterances' here may seem surprising, but semantics must presumably be involved in utterance interpretation: pragmatics needs linguistic content to operate on.

RATIONALITY, PRAGMATICS, AND SOURCES

1. Make your contribution as informative as required (for the current purposes of the exchange)
2. Do not make your contribution more informative than required.

Quality: Super-maxim, “Try to make your contribution one that is true”

1. Do not say what you believe to be false
2. Do not say that for which you lack adequate evidence

Relation: Be relevant

Manner: Super-maxim, “Be perspicuous”

1. Avoid obscurity of expression
2. Avoid ambiguity
3. Be brief (avoid unnecessary prolixity)
4. Be orderly

(Grice, 1975, pp. 45-46)

A speaker can choose to observe or ostentatiously flout the maxims to generate an implicit proposition, or implicature. For example:

Maxim Observed

A: I’ve just run out of petrol.

B: Oh, there’s a garage just around the corner.

Implicature: A may obtain petrol there.

Maxim Flouted

A: Let’s get the kids something to eat.

B: Okay, but I veto I-C-E-C-R-E-A-M.

Implicature: B would rather not mention ice cream directly in case the children then demand some.

(Levinson, 1983, p. 104).

Although Grice’s (1975) framework is suggestive and influential, it is imperfect. The imperfections suggest a more minimal, psychologically plausible account. Grice himself does not appear to have viewed his account as definitive (B.

RATIONALITY, PRAGMATICS, AND SOURCES

Clark, 2013) or psychological (Noveck & Reboul, 2008). His framework can account well for qualitative data (for discussion, see Huang, 2007; B. Clark, 2013); but it does so at the expense of some vagueness (B. Clark, 2013). For instance, Grice does not define relevance here. Even so, pragmatic analyses tend to cite the Maxim of Relation or the Maxim of Quantity to the exclusion of the other maxims, giving rise to accusations of redundancy (B. Clark, 2013). Consequently, prominent successors to Grice have reduced his framework by fleshing out the notion of relevance (Sperber & Wilson, 1995) or quantity (Horn, 1984, 1989, 2004). While the original Gricean framework has inspired psychological research (e.g. Surian, Baron-Cohen, & Van der Lely, 1996), more minimal, more psychologically specific successors have proved more stimulating (e.g. Breheny, Ferguson, & Katsos, 2013; Breheny, Katsos, & Williams, 2006; Geurts, Katsos, Cummins, Moons, & Noordman, 2010; Noveck & Reboul, 2008; Noveck & Sperber, 2004). There are numerous other theories in pragmatics (e.g. Levinson, 2000) and semantics (e.g. Jaszczolt, 2007; Kempson, Meyer-Viol, & Gabbay, 2000), but these also tend to be more minimal than Grice's framework. For present purposes, we do not need to decide on a pragmatic theory; it suffices simply to note the kind of mentalistic assumptions that underpin pragmatics.

While pragmatic theories have become more streamlined, the purview of pragmatics has increased considerably. Pragmatics intrudes still further into the interpretation of any utterance, including experimental materials. For Grice, semantics contributed what is said⁸ - the conventional content of an utterance – and pragmatics largely contributed what is implicated (B. Clark, 2013). In other words,

although pragmatics could leak into conventional content, through reference assignment for indexical expressions, such as pronouns, and the resolution of ambiguity, pragmatics' primary business was calculating implicatures⁹ (see, e.g., B. Clark, 2013; Grice, 1975). Subsequently, researchers have argued that pragmatics contributes far more profoundly to what is said. For instance, a commonplace word such as 'fresh' can have a great many meanings, as witnessed by collocations such as 'fresh shirt', 'fresh vegetables', 'fresh bread', 'fresh air', and 'fresh idea' (Wilson & Carston, 2007). One response is that words are radically ambiguous and require heavy pragmatic work to disambiguate them; another, perhaps more plausible response is that word meanings are adjusted *ad hoc* (Wilson & Carston, 2007). This issue will receive a more detailed treatment in Chapter 2, where potential effects on reasoning are discussed.

As we have seen, pragmatics connects with sources (speakers) through theory of mind, at least in the relative shallow sense of interpreting utterances using somewhat generalized assumptions about speakers' behaviour (e.g. conversational maxims). Deeper social knowledge is also implicated, drawing pragmatics and sources still closer together. This social knowledge can take the form of facts about social institutions such as promises, marriages, courtroom behaviour, and so on (see, e.g., Austin, 1965; for discussion, see Korta & Perry, 2015). More importantly, for present purposes, recent theories suggest that, when hearers interpret an utterance, they also refer to deeper, more stable attributes of the speaker. Hilton (1995) suggests that Gricean maxims are linked to character traits, providing information in

⁸ For theoretical discussion of 'what is said', see Carston (2008) or Recanati (2002, 2004).

both directions. For example, respecting the maxims makes a speaker appear cooperative; and knowledge about the speaker's character can help the hearer decide whether a speaker is respecting a maxim. Hilton gives the example of an underwhelming reference for a job. Should we infer that the reference is meant as a non-recommendation? This inference will likely be strengthened if we know that the referee is usually helpful and generous with praise. But the referee's pragmatic behaviour may, in turn, affect our assessment of their cooperativeness, likeability, and so on.

More recently still, two theories have proposed a strong link between pragmatics and source reliability. This link will be especially prominent in subsequent chapters. Sperber et al. (2010) argue that pragmatics is supported by a tendency towards epistemic vigilance. Hearers take an initially trusting stance, being prepared to update their beliefs in response to speakers' statements; but they monitor speakers' trustworthiness, and the content of messages, to guard against deception and misinformation. In formal linguistics, McCready (2014) makes a similar case, arguing that pragmatics depends on cooperation, which in turn depends on trust. On his view, speakers are careful to manage their reputations, to maintain sufficient levels of cooperation. If speakers' reputations and trust drop low enough, then cooperation – hence, pragmatics – ceases.

These considerations suggest an important link between pragmatics and the stable attributes of a speaker: in the terminology of this thesis, between pragmatics and sources. This link is underexplored, with one notable exception. Empirical work

⁹ We can understand 'implicatures', here, to include both metaphors and irony, though see, for example, Wilson and Carston (2007) for an alternative pragmatic account of metaphors.

RATIONALITY, PRAGMATICS, AND SOURCES

has tended to focus on shallow theory of mind, in which intended meanings are inferred using rather generalized assumptions about speakers' behaviour. A different approach is taken by Harris, Corner, and Hahn (2013). They explored the phenomenon of 'damning with faint praise', taking the example, familiar from above, of an underwhelming reference. If a reference describes an applicant for a mathematics course only as 'polite and punctual', then it would certainly be underwhelming: so underwhelming, in fact, that it could be taken as a warning not to admit the applicant. Harris, Corner, and Hahn formalized this example as an argument from ignorance: since there is no evidence to show that the referee thinks that the applicant is a good mathematician, we conclude that he is not. Harris, Corner, and Hahn predicted that the strength of this inference depends on the expertise of the source (referee). They found supporting experimental evidence. Participants read a recommendation letter for a maths course, and rated how strongly they believed that the candidate should be admitted. Ratings decreased for the underwhelming 'polite and punctual' reference but only when the source was the applicant's maths tutor, not when it was merely the applicant's personal tutor. Here, source expertise determined whether participants inferred that the reference was a non-recommendation and decreased their ratings.¹⁰

A picture emerges, then, of a pragmatics that is social, mentalistic and inferential, and that is fundamental to interpreting utterances, be they utterances in natural conversations or in psychological experiments. Pragmatics, on this account, is a key component in Crupi's framework for assessing rationality in psychological

¹⁰ Note that Harris, Corner, and Hahn also explicitly represent a pragmatic inference as an argument, an approach which assumes a close link between pragmatics and argumentation.

studies. Pragmatics imposes ‘evidential requirements’ on such studies (Lee, 2006, p. 194): it is crucial to ‘[get] the communication right’ before proclaiming irrationality (Schwarz, 1996, p. 5). This is not to say that pragmatics will allow us to explain away all apparent irrationality, but rather that ‘we need to ensure that our experiments do not produce the phenomena for reasons that are unlikely to hold in daily life’ (Schwarz, 1996). In other words, pragmatics allows us to test the robustness of the evidence for irrational behaviour.

1.2.2 Pragmatics and Theories of Rationality

Now that we have a clearer picture of pragmatics, we can consider how it features in the theories of rationality discussed above. This process is not straightforward, because the theories adopt different and, in some cases, remarkably vague definitions of pragmatics.

Perhaps the closest to the present definition is Oaksford and Chater’s approach (2007, 2009), which invokes pragmatics to explain a mismatch between their model and data on conditional inference. Although Oaksford and Chater seem to take a broadly Gricean view of pragmatics, their Bayesian Rationality also meshes well with new Bayesian approaches which promise rich treatments of pragmatic phenomena (Frank & Goodman, 2012; Franke & Jäger, 2016; Goodman & Lassiter, in press; Goodman & Stuhlmüller, 2013). The link between Bayesian rational and psycholinguistic models needs further exploration.

Other theories invoke pragmatics to explain more fundamental mismatches between basic theoretical predictions and the data. For instance, both Mental Logic and Mental Models assume some kind of Gricean (or post-Gricean) pragmatics. Both also accommodate knowledge about social situations (Braine & O’Brien, 1991; Johnson-Laird & Byrne, 2002), perhaps embodied in pragmatic-reasoning schemas

RATIONALITY, PRAGMATICS, AND SOURCES

(Braine & O'Brien, 1991). But since their pragmatics includes broader general knowledge, it is a pragmatics so broad that it is hard to imagine a general, predictive account. For instance, when Mental Models Theory is implemented computationally, pragmatics, in the form of general knowledge, is simply written into the computational model for a specific context (Johnson-Laird & Byrne, 2002). For example, a computational model will specify that wet matches will not light, to cancel an inference from 'A match is struck' to 'The match will light'. Although this strategy is successful in individual cases, it is *ad hoc* and unlikely to generalize.

Still other theories invoke pragmatics to explain why apparent irrationality occurs. In dual-process theories, pragmatics is typically labelled a Type 1 process (Stanovich & West, 2000), a decision which is surprising given the ample evidence that pragmatic inferences are effortful and can be disabled by cognitive load (Noveck & Reboul, 2008) and resemble rational arguments (Macagno & Walton, 2013). Nevertheless, in dual-process theory, pragmatics is a cause of biased reasoning. Lastly, the Fast and Frugal Heuristics program uses pragmatics to explain *away* irrationality: pragmatics is part of social, hence ecological, rationality, and is one way to reduce uncertainty (Gigerenzer, 2010, 2015)¹¹.

This thesis takes a different strategy from the theories above. It does not invoke pragmatics to explain away problematic data for a given theory or position on rationality; rather, it argues that pragmatics and rationality are fundamentally –

¹¹ Missing from this discussion is the approach of Stenning and van Lambalgen (2012). It is unclear how pragmatics fits with their account. Stenning and van Lambalgen posit specialist logics for mentalistic inference. Pragmatics could correspond to one such logic. But pragmatics is also relevant for what they call credulous reasoning - finding a single interpretation that makes the speaker's utterances true – which sounds decidedly pragmatic. Lastly, a crucial notion is reasoning to an interpretation: selecting the appropriate logic for the task in hand. This reasoning, too, could be considered pragmatic.

indeed, inextricably – linked. The thesis follows Crupi (personal communication, September 5, 2016) in arguing that considering pragmatics is a precondition for establishing irrationality in an experiment, and it follows Schwarz (1996) and Lee (2006) in arguing that pragmatics tests the robustness of apparent irrationality. But it will argue, in later chapters, that there is a deeper connection between pragmatics and rationality that justifies treating the two together.

The next section will discuss empirical studies that use pragmatics to test the robustness of apparent irrationality. The aim is to present, not an exhaustive survey, but a sample of diverse studies which show how pragmatics has tended to feature in the literature on (ir-)rationality.

1.2.3 Pragmatics and Experiments on Rationality

The psychologists Denis Hilton (1995) and Norbert Schwarz (1996) have done much to emphasize the importance of considering pragmatics when assessing rationality. Experiments, they observed, are communicative situations of a peculiar kind. Experimental tasks differ substantially from communication in daily life. For reasons of experimental control, instructions, questions and response options have standardized wordings; and, if asked, experimenters will generally not help to clarify their meanings, for fear of introducing confounds (Schwarz, 1996). In such contexts, participants may rely even more on general pragmatic expectations than they would in daily life (Schwarz, 1996). Given that experimenters will often covertly behave in non-cooperative ways – for instance, by providing deliberately misleading or irrelevant information (Hilton, 1995; Schwarz, 1996) – pragmatics has the potential to explain much of the variance in experimental tasks. Only once this variance is accounted for can we pronounce on rationality. This section will review major

findings from attempts to address this issue, and will relate them to Crupi's framework: more specifically, to premise (mis-)match and conclusion (mis-)match.¹²

Premise Mismatch.

Presupposition. A speaker often takes for granted, or presupposes, certain information when making an utterance. Imagine, for instance, a speaker making the request, 'Please put the bins out when you leave for work this morning.' The speaker, here, presupposes that the hearer is going to work on the morning in question. Presupposition features in a cluster of related effects which, at first blush, suggest that people naively accept presupposed information.

P1. 'Did you see the children getting on the school bus?'

Asked this question after seeing a video, participants are more likely than controls to falsely remember seeing a school bus even though there was none (see, e.g., Loftus, 1975).

P2. 'What do you do to liven things up at a party?'

Participants form the impression that the speaker is an extravert, on scant information (see, e.g. Swann, Giuliano, & Wegener, 1982).

P3. 'The CIA is not currently involved in illegal drug trafficking.'

After reading this denial in a fictional newspaper report, participants who had indicated prior disbelief in this claim nevertheless increased their belief that the CIA is, in fact, involved in illegal drug trafficking (Gruenfeld & Weyer, 1992).

¹² At no point will I assume that the reviewed studies provide the definitive analyses of the tasks in question. These studies simply show that pragmatic factors must be considered.

RATIONALITY, PRAGMATICS, AND SOURCES

Participants seem to be accepting presupposed information without recognizing a possible difference between the speaker's belief and reality. The false-memory effect, as in P1, has been taken as a sign of irrationality (for discussion, see Hilton, 1995, and Schwarz, 1996). But a pragmatic analysis, supported by the effects in P2 and P3, suggests a counterargument. This counterargument relies on recognizing the presupposed information. In P1 and P2, the speaker's presuppositions are relatively clear: that there was a school bus, and that parties should be lively. In P3 the presupposition is subtler: that there were enough grounds for suspicion to issue a denial.

We turn to the pragmatic analysis. False-memory effects, as in P1, seem to require the leading question to come from a trustworthy (Dodd & Bradshaw, 1980) or knowledgeable source (Smith & Ellsworth, 1987)(for discussion, see Hilton, 1995; Schwarz, 1996). Thus, people may be rationally changing their beliefs because of the testimony of reliable sources. Admittedly, the false-memory effect remains surprising, at least to the extent that people are unaware of, or do not report, this inference, and to the extent that they experience memories of events whose occurrence they take on trust. However, the effects in P2 and P3 offer supportive evidence that people handle presupposed information critically. Impression formation, as in P2, relies on a cooperative speaker, and does not survive when the information is described as randomly drawn (Swann, Giuliano, & Wegner, 1982; for discussion, see Schwarz, 1996). Furthermore, 'saying things that go without saying'¹³, as in P3, seems to rely on a pragmatic context. For instance, to borrow Gruenfeld & Wyer's (1992) example, imagine that we are told that 'Peter's IQ is at

¹³ This is Schwarz's (1996) term for the effect.

least 100'. If a hearer lacks any knowledge of Peter's IQ, then this utterance could be taken at face value: Peter is of at least average, and perhaps above average, intelligence. However, if the hearer had previously believed Peter to be his high-school valedictorian, then, to avoid apparent irrelevance, this utterance could be taken to mean that Peter is only of average, not above average, intelligence. Thus, in the CIA case above (and equivalent materials), when the source was a reference source (e.g. an encyclopaedia), the information could be taken at face value, and belief in the denied proposition did not significantly increase. But when the source was a newspaper, taken to be conveying novel, newsworthy information, then belief in the denied proposition increased significantly (Gruenfeld and Wyer, 1992).

This nuanced view of presupposition contrasts with the findings of Gilbert and colleagues (D. T. Gilbert, Krull, & Malone, 1990; D. T. Gilbert, Tafarodi, & Malone, 1993) which suggest that, at base, people naively believe what they are told. But, again, a pragmatic defence is possible. In probably the most widely discussed experiment, participants read definitions of words purportedly from the Hopi language (though actually non-words), such as 'A monishna is a star'. The definitions were followed by the tag 'true' or 'false'. Afterwards, participants had to perform a memory test. On some trials, processing was disrupted, for instance, with a tone. Disruption significantly increased the number of false trials remember as true, suggesting that, by default, people represent new information as true, and that scepticism takes extra work and is fragile. However, Hasson, Simmons, and Todorov (2005) argued that the experiment confounded truth value and informativeness¹⁴.

¹⁴ A relevance theorist (e.g. Sperber & Wilson, 1995) would likely want to re-express this analysis in terms of relevance: more specifically, here, in terms of cognitive effects. This would be a straightforward step, but is not necessary for present purposes.

RATIONALITY, PRAGMATICS, AND SOURCES

Negative utterances such as ‘A monishna is not a star’ are, in this context, radically uninformative, as ‘a monishna’ could refer to anything else. Hasson, Simmons and Todorov found that, when sentences were informative when false (e.g. ‘This person owns a television’), or when true or false (‘This person is a liberal’), the effect disappeared. This finding suggests that people can adopt a somewhat critical stance to new information (see, also, Richter, Schroeder, & Wöhrmann, 2009).

Relevance. As we saw above, the notion of relevance is central to several major theories of pragmatics¹⁵. Relevance has featured, too, in the pragmatic analyses of apparent irrationality. A good example is the literature on base-rate neglect. Kahneman and Tversky (1973) identified this effect in a simple prediction task. Participants read a scenario concerning a group of engineers and lawyers: for instance, in one version, 30 engineers and 70 lawyers. They were asked to predict whether a target person was an engineer or a lawyer. Target people were introduced with descriptions such as the following:

Jack is a 45-year-old man. He is married and has four children. He is generally conservative, careful, and ambitious. He shows no interest in political and social issues and spends most of his free time on his many hobbies, which include home carpentry, sailing, and mathematical puzzles.

(Kahneman & Tversky, 1973, p. 241)

This description is meant to be representative of an engineer. Participants tend to rely on such descriptions when making their judgments and to neglect the provided

¹⁵ As we have also seen, an alternative is to focus on the maxim of quantity. It may be possible to translate between the notions of quantity and relevance to some extent (for discussion, see Huang, 2007). This theoretical project will not be undertaken here.

RATIONALITY, PRAGMATICS, AND SOURCES

base rates. Thus, in this case they would tend to ignore the stronger prior probability of Jack being a lawyer (.7) and predict that he is an engineer.

This apparent base-rate neglect is striking, but the task is pragmatically complex. The task presents irrelevant information - the psychological descriptions – and emphasizes its quality by identifying its source as expert psychologists (Schwarz, 1996). When the information is attributed to a computer's random selection, participants can produce predictions much closer to the base rate (Schwarz, Strack, Hilton, & Naderer, 1991). The classic task also varies the descriptions within-subjects but not the base rates, making the descriptions seem highly relevant (Schwarz, 1996). Varying the base rates across trials encourages participants to make greater use of them (Schwarz, Strack, Hilton, & Naderer, 1991). Importantly, although such manipulations reduce base-rate neglect, participants still tend to underuse base rates (Schwarz, Strack, Hilton, & Naderer, 1991). The important point, here, is that pragmatic features of Kahneman and Tversky's (1973) original task led to overestimating the size and robustness of the effect.

A similar case can be made for the effect known as the fundamental attribution error. This effect arises when participants are asked to explain a person's behaviour: they tend to over-weight the person's characteristics (e.g. their intentions) and underweight the situation (for discussion, see Schwarz, 1996). In a typical task, participants read about students who have been told to write essays arguing for a position that has been chosen for them (Schwarz, 1996). Even though participants are aware of this strong situational factor, they nevertheless tend to attribute the position to the students. However, the experimental design echoes that of the base-rate neglect study in that it deliberately presents irrelevant information. The typical instructions and materials may not be clear enough to convince participants to reject

RATIONALITY, PRAGMATICS, AND SOURCES

potentially relevant information: in this case, the essay content (Schwarz, 1996). In a clearer version of the task, Wright and Wells (1988) instructed participants that the information about students was randomly selected and potentially incomplete. In this case, the size of the effect was approximately halved, though the effect was not eliminated.

In the studies cited above, relevance is understood in an intuitive, informal way. Relevance is put to work in a more theoretical way in an analysis of the Wason selection task (for the original task, see, Wason, 1966). Sperber, Cara, and Girotto (1995) used Relevance Theory (e.g. Sperber & Wilson, 1995) to explain participants' performance on this staple of the psychology of reasoning. The standard abstract version of the task requires participants to test a rule by turning over some cards. Consider, for example, the following sample item as presented by Van der Henst and Sperber (2004):

Rule: If there is a 6 on one of the cards, there is an E on the other.

Card 1: 6 Card 2: 7 Card 3: E Card 4: G

Formal logic prescribes that, to falsify a rule of the form 'If P, then Q', we must find an instance of 'P and not Q' and that, therefore, we must select the cards '6' and 'G' above. In general, only around 10% of participants select these cards (Sperber, Cara, & Girotto, 1995; Van der Henst & Sperber, 2004). Although other, non-logical analyses are possible (see, e.g., Oaksford & Chater, 1994), Sperber, Cara, and Girotto (1995) stick to formal logic. They argue that Relevance Theory can explain the standard response, and can successfully predict when people will give the supposedly normative response.

On this Relevance account, participants interpret the task as an act of communication, and deploy a relevance-based inferential mechanism in the

RATIONALITY, PRAGMATICS, AND SOURCES

following way. Relevance, here, is understood as a trade-off between cognitive effects (e.g. true conclusions, or strengthened/weakened assumptions) and processing effort (Sperber & Wilson, 1995; Sperber, Cara, & Girotto, 1995). The more cognitive effects, the more relevant the stimulus; the higher the processing effort, the less relevant the stimulus. Cognitive effects are derived following a path of least effort. Sperber, Cara, and Girotto (1995) argue that, from a general conditional rule, two inferences are particularly accessible. Firstly, *modus ponens*: participants infer that the '6' card must have an E on the other side, and therefore select it. Secondly, the rule has instances: participants infer that there are cards that match the rule, and therefore select the '6' card and the 'E' card. These are the two most common responses to the standard abstract task (Sperber, Cara, & Girotto, 1995). It requires, in contrast, high processing effort to infer from the given rule 'Not (6 and not E)'.

The above analysis makes the following prediction: manipulating cognitive effects and processing effort should lead to more logical responses. More specifically, cognitive effects can be boosted by, for instance, making 'P and not Q' cases more salient; and processing effort can be minimized by, for instance, using lexicalized 'P and not Q' cases (as, for example, bachelors are male and not married). Sperber, Cara, and Girotto (1995) manipulated these parameters in a set of experiments and found support for their predictions. For instance, in one task, they used the rule 'If a woman has a child, then she has had sex'. On their view, a 'not Q' case here would have considerable cognitive effects but would be highly implausible. The authors created a scenario which made 'not Q' plausible: the leader of a cult was rumoured to be forcing female members to undergo artificial insemination to create a generation of 'Virgin-Mothers'. Here, 'not Q' is plausible but not certain. Moreover,

RATIONALITY, PRAGMATICS, AND SOURCES

participants were told to imagine that they were journalists writing about the cult. In this case, 78% of participants selected the ‘P and not Q’ cards (Virgin-Mothers). This method has been extended to the deontic version of the selection task with equivalent results (Giroto, Kimmelmeyer, Sperber, & van der Henst, 2001).

Once again, pragmatics can explain important patterns of participant behaviour. Here, however, pragmatics plays a different role from that in the other reviewed studies. While the other studies argued for new estimates of the size and prevalence of existing effects, Sperber, Cara and Giroto (1995) argued that the Wason Selection Task is not fit for purpose. For, on their view, participants do not engage in reflective reasoning at all, but rather in spontaneous communicative inferences. In other words, Sperber, Cara and Giroto take the literature on the Wason Selection Task to be irrelevant to the rationality debate. This point of view is by no means universally accepted (see, e.g., Manktelow, 2012; Stenning & van Lambalgen, 2008, for discussion).

Framing and Implicature. One of the key phenomena in judgment and decision making is the framing effect (for a review, see Levin, Schneider, & Gaeth, 1998). A framing effect occurs when different representations of a problem prompt different responses. Simple examples arise often in consumer-judgment research. For instance, people tend to judge the same product differently when it is described as 75% lean or 25% fat (Levin, Schneider, & Gaeth, 1998). This phenomenon is widely taken to be problematic, on the grounds that the information is objectively or ‘informationally’ equivalent (for discussion, see Corner & Hahn, 2010; Levin et al., 1998). Pragmatics, however, offers a different perspective (Corner & Hahn, 2010; McKenzie & Nelson, 2003; Sher & McKenzie, 2006).

We will start with the simplest type of framing, attribute framing, where framing is applied to a single underlying attribute or variable (Levin, Schneider, & Gaeth, 1998). The following are examples of attribute frames: describing a product as $N\%$ fat-free/ $100-N\%$ fat; describing an operation as having an $N\%$ survival/ $100-N\%$ mortality rate; describing a business team as having an $N\%$ success/ $100-N\%$ failure rate. The standard finding is that people give higher favourability ratings to an item when it is described using a positive frame (e.g. $N\%$ fat-free) than when it is described using a negative frame (e.g. $100-N\%$ 'fat')(Levin, Schneider, & Gaeth, 1998). McKenzie and colleagues (McKenzie & Nelson, 2003; Sher & McKenzie, 2006) disputed the conventional assumption that positive and negative frames convey equivalent information, and argued that information leaks through the supposedly equivalent descriptions¹⁶. On their view, we understand such situations relative to an implicit reference point: for instance, the expected, or normal, level of fat/fat-freeness. All else equal, speakers are more likely to select 'fat' (respectively, 'fat-free') if there is an increase relative to the reference point, that is, there is more fat (respectively, 'fat-freeness') than usual. Listeners are sensitive to these regularities, and draw appropriate inferences (or implicatures).

McKenzie and colleagues (McKenzie & Nelson, 2003; Sher & McKenzie, 2006) explored this account in a series of experiments. In one task, for instance, they asked participants to imagine that they had a 4-ounce measuring cup in front of them. One group of participants was told to imagine that the cup was filled with water to the 4-ounce line; that they left the room; and that on their return they found that the

¹⁶ Sher and McKenzie (2014) generalized this into an 'options as information' account, in which choice sets are taken to leak information, accounting for preference reversals.

RATIONALITY, PRAGMATICS, AND SOURCES

water was at the 2-ounce line. Another group of participants was told to imagine the scenario but with the cup initially empty and then, after their return to the room, with water at the 2-ounce line. In each case, participants chose whether to describe the cup as half-full or half-empty. Participants were significantly more likely to describe the cup as half-full when it was previously empty, and as half-empty when it was previously full. These data support the reference-point hypothesis. In a separate task, participants were presented with a different scenario in which a participant described a cup as ‘half-full’ or ‘half-empty’, and were asked to decide whether the cup was previously full or previously empty. Participants were significantly more likely to choose ‘previously full’ when the cup was described as ‘half-empty’, and ‘previously empty’ when the cup was described as ‘half-full’. These data also support the reference-point hypothesis (see, also, Ingram, Hand, & Moxey, 2014, and Keren, 2007; for a related semantic analysis, see Geurts, 2013).

A pragmatic analysis has also been offered for another type of framing, risky-choice framing: in particular, the so-called Asian Disease Paradigm. In this paradigm, participants see the following types of material:

Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate[s] of the consequences of the programs are as follows:

[Usually seen by one group]

If Program A is adopted, 200 people will be saved.

If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.

RATIONALITY, PRAGMATICS, AND SOURCES

[Usually seen by another group]

If Program C is adopted, 400 people will die.

If Program D is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.

(Tversky & Kahneman, 1981, p. 453)

Note that Programs A and C are low risk (risk-averse) options; Programs B and D are high risk (risk-seeking) options. Participants who see the pair of Programs A and B tend to prefer low risk Program A (72% in Tversky & Kahneman, 1981). Those who see the pair of Programs C and D tend to prefer high risk Program D (78% in Tversky & Kahneman, 1981). The materials manipulate framing through the descriptions ‘will be saved’ and ‘will die’. On the standard account (e.g. Tversky & Kahneman, 1981), the phrase ‘will be saved’ fixes the reference point at 0 lives saved (600 lives lost). Participants treat any survivors, then, as gains relative to the reference point, and are risk-averse. In contrast, the phrase ‘will die’ fixes the reference point at 0 lives lost (600 lives saved). Participants treat any victims, then, as losses relative to the reference point, and are risk-seeking. Generally, risky-choice framing replicates well with these, and equivalent, materials (for a review, see Levin, Schneider & Gaeth, 1998; for more recent replications of the basic effect, see Costa, Foucart, Arnon, Aparici, & Apesteguia, 2014, and Keysar, Hayakawa, & An, 2012)

This species of framing has, however, provoked controversy: the options, and framing conditions, differ in information content (Kühberger, 1995; Kühberger & Tanner, 2010; Mandel, 2001, 2014). Programs A and C specify only what happens to 200 or 400 people respectively; Programs B and D fully specify the outcomes. This pattern of information may be a prerequisite for the classic framing effect. The framing effect can be blocked by fully specifying all programs (e.g. Kühberger,

RATIONALITY, PRAGMATICS, AND SOURCES

1995; Mandel, 2001). It can be reversed by specifying only the complements: for example, in Program A, ‘400 people will not be saved’ (Kühberger, 1995; see, also, Kühberger & Tanner, 2010). People do not seem to use arithmetic to infer what happens to the unmentioned people. Such arithmetic would presuppose that participants interpret the number terms to mean ‘exactly N people’. In fact, qualifying number terms with ‘exactly’ can render framing non-significant; conversely, qualifying them with ‘at least’ can increase framing (Mandel, 2014). Participants also report interpreting the unmodified number terms as ‘at least’ (Mandel, 2014). For such participants, Program A has a potentially better outcome than B; Program D, a potentially better outcome than C.

Although such findings are highly compelling, there is reason, still, to believe in risky-choice framing. For one thing, the Asian Disease Paradigm only corresponds to a subset of risky-choice data (Tversky & Kahneman, 1981). For another, the Asian Disease Paradigm may be over-determined. In a recent study, Chick, Reyna, and Corbin (2016) tested the robustness of framing in the Asian Disease Paradigm while controlling for the ‘at least’ reading of the number term. They provided extensive instructions, detailed examples, and quizzes, and ran various analyses, including focusing on participants who reported an ‘exactly’ reading of the number terms. This method yielded robust framing effects. As the authors observe, ‘at least’ readings may be a sufficient condition for a framing effect but do not appear to be a necessary one. More data are needed to explore how robust such framing is to pragmatic manipulations (for an alternative approach, see Wallin, Paradis, & Katsikopoulos, 2016).

Another type of framing, goal framing, will arise in a later chapter. For now, we will turn to a phenomenon that bridges premise (mis-)match and conclusion (mis-)match.

The Conjunction Fallacy: Bridging Premise (Mis)match and Conclusion (Mis)match.

The conjunction fallacy is one of the most widely debated effects in the judgment literature. Probably the simplest version of the effect is the following, from Tversky and Kahneman (1983, p. 299)

Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Linda is a bank teller. (T)

Linda is a bank teller and a feminist. (T&F)

Participants were asked to indicate which of (T) and (T&F) was the most probable. 85% chose (T & F), violating the conjunction rule of probability: here, $P(T\&F) \leq P(T)$ (Tversky & Kahneman, 1983). Tversky and Kahneman (1983) demonstrated the effect with different response formats and materials, and a large literature has manipulated various parameters (for a review, see Moro, 2009).

The conjunction fallacy has prompted numerous pragmatic analyses. We begin with premise (mis-)match. Firstly, participants may understand the term 'probability' differently from the mathematical sense intended by the experimenters. They may understand it as, for instance, plausibility, believability or imaginability (Fiedler, 1988; Gigerenzer, 1996; Hertwig & Gigerenzer, 1999). There is some merit to this suggestion: betting versions, which avoid talking of probability, tend to

RATIONALITY, PRAGMATICS, AND SOURCES

produce lower fallacy rates, but nevertheless over 50% (Bar-Hillel & Neter, 1993; Bonini, Tentori, & Osherson, 2004; Messer & Griggs, 1993; Sides, Osherson, Bonini, & Viale, 2002; Tversky & Kahneman, 1983; Wolford, Taylor, & Beck, 1990).

Likewise, a frequency format reduces, but does not eliminate, the fallacy (Fielder, 1988). Secondly, participants may take their task to be identifying the most informative response (Hertwig & Gigerenzer, 1999). But there does not appear to be empirical evidence to support this interpretation (Moro, 2009).

We turn to conclusion (mis-)match: to different interpretations of the response set. Firstly, people may not interpret the linguistic conjunction as a logical/probabilistic conjunction, since the linguistic conjunction is ambiguous. The most relevant alternative readings are disjunction and causation (for discussion, see Moro, 2009). However, the conjunction fallacy occurs even when ‘and’ is not used in the materials (Bar-Hillel & Neter, 1993). And it occurs, albeit at a reduced rate, when controlling for disjunctive readings (Bonini, Tentori, & Osherson, 2004; Sides, Osherson, Bonini, & Viale, 2002). Moreover, causal readings seem to be interpreted not as conditionals (*contra* Hertwig, Benz, and Krauss, 2008), but as conjunctions, to which the conjunction rule still applies; and these causal conjunctions, then, seem to produce a robust conjunction fallacy (Tentori and Crupi, 2012). Nevertheless, some manipulations do seem to substantially reduce the fallacy: Politzer and Noveck (1991) found that implicit conjunctions, such as ‘Lendl will win the finals’ (as opposed to ‘Lendl will play the finals and win’) significantly decreased the occurrence of the fallacy (though see Moro, 2009, for a critical discussion).

Secondly, people may interpret the individual conjuncts differently: for instance, they might interpret ‘Linda is a bank teller’ to implicitly mean ‘and not a feminist’ (Tversky & Kahneman, 1983). Researchers have responded principally by

exploring different wordings and altering the response set (Moro, 2009). A range of re-wordings of ‘and’ have prompted different fallacy rates: ‘whether or not’ – 57% (Tversky and Kahneman, 1983); ‘regardless of whether or not’ – 56% (Messer & Griggs, 1993); ‘may or may not’ – 69%¹⁷ (Agnoli & Krantz, 1989). A complex re-wording by Dulany and Hilton (1991) reduced the rate to 38%, but contains other cues for normative behaviour (Moro, 2009). Changes to the response set have also varied the fallacy rate. Including both conjuncts, (T) and (F), reduced the rate in an estimation task (42%) but not a ranking task (78%) (Hertwig & Chase, 1998). Adding a disjunction of the events, (T v F), somewhat reduced the rate to 69% (Morier & Borgida, 1984). Perhaps most strikingly, however, including items equivalent to (T), (F), (T&F), (T & not F) produced a high rate, 78%, even in a betting task which should prompt more probabilistic reasoning (Bonini et al., 2004).

Overall, there is evidence that pragmatic factors influence the robustness of the conjunction fallacy: both premise (mis-)match and conclusion (mis-match). Given the range of these factors, it is hard to get a general sense of the robustness. Nevertheless, the conjunction fallacy does seem to survive.

Conclusion (Mis)match.

Other examples of conclusion (mis-)match abound in survey research. Here, survey responses may not be literal. For instance, many survey participants will give answers to questions about fictional issues (for discussion, see Schwarz, 1996). A straightforward interpretation is that participants dislike appearing ignorant and, therefore, make up a response as a face-saving device (Schwarz, 1996). This possibility raises alarming questions about the quality of survey research. However,

¹⁷ Macdonald and Gilhooly (1990) found 20%, switching the problem to the future.

RATIONALITY, PRAGMATICS, AND SOURCES

a pragmatic interpretation is plausible. Strack, Schwarz, and Wänke (1991) hypothesized that participants use context – for instance, surrounding questions – to transform the question into one about a broader set of issues on which they can legitimately offer an opinion. They asked student participants about a fictional ‘educational contribution’, and used a previous question to suggest that the contribution meant the state paying the student or the student paying the state. Participants’ views depended on this previous question. It remains to be seen how often this pragmatic account is correct.

Survey and experimental research alike make use of response scales. Varying these scales can cause considerable inconsistencies in data, suggesting unreasonable responding (Schwarz, 1996). However, scale wording can often be vague or ambiguous, and scale numbers can cue specific interpretations. Take, for instance, a question about how successful a participant is in life, with a scale anchored at the lower end with ‘not at all successful’ or ‘unsuccessful’. These terms are ambiguous. If the numeric scale ranges from -5 to 5, it prompts participants to think of failure; if the scale ranges, instead, from 0 to 10, it prompts them to think of the absence of success (Schwarz, Knauper, Hippler, Noelle-Neumann, & Clark, 1991). Participants’ responses cluster significantly higher when the -5 to 5 scale is used. Similarly, Schwarz (1996) reports data that frequency scales, lower-bounded at ‘rarely’, receive different interpretations depending on the numeric scale: if the scale is anchored at 0, participants interpret ‘rarely’ as ‘never’; if at 1, as ‘sometimes’. Such frequency scales can also act as cues for estimating the frequency of habitual events, such as watching television, which may not be represented discretely in memory (Schwarz, 1996). Lastly, these scale phenomena can interact with the interpretation of questions. For instance, a vague question like ‘How often are you really irritated?’ can be

interpreted differently depending on the scale. If the frequency scale is anchored at ‘several times daily’ and ‘less than once a week’, participants report minor irritations; if the scale is anchored, instead, at ‘several times a month’ to ‘less than once every three months’, participants report more major irritations (Schwarz, 1996).

Summary.

This section has argued that pragmatics is essential to the rationality debate. The grounds for this are multiple. Firstly, pragmatics features, to a varying degree, in prominent psychological theories. Secondly, pragmatic analysis is a precondition for assessing rationality in experiments: we must check whether participants and experimenters understand the premises (instructions and materials) and conclusions (responses) in the same way. Considering pragmatics in this way is both theoretically essential and richly productive. A large and growing literature has emerged, testing the size and robustness of common effects, principally biases in decision making and judgment. This focus on biases should not suggest, however, that pragmatics is only relevant to diagnoses of irrationality. When people show apparently rational behaviour, such effects, too, should be pragmatically analysed, to ensure that they are not merely an artefact of a particular design. Finally, later chapters will argue for a stronger conceptual relationship between rationality and pragmatics.

1.3 Sources

We turn now to the third component in the triangular scheme, sources. This thesis will argue that sources are inextricably linked to both rationality and pragmatics. The link to pragmatics has already been explored above. This section will introduce the link between sources and rationality.

Within psychology, sources have been given the most extensive treatment in the psychology of persuasion and, in particular, in dual-route models. Although this

RATIONALITY, PRAGMATICS, AND SOURCES

social psychological research is relevant for debates on rationality, it takes a purely descriptive approach. The persuasion literature will be surveyed in more detail in Chapter 6. Here, it suffices to note that information about sources – typically, a person making a claim - generally features in the peripheral, heuristic route (Briñol & Petty, 2009; Chaiken, 1980; Chaiken & Maheswaran, 1994; Petty & Briñol, 2008; Petty & Cacioppo, 1984, 1986). That is, sources influence persuasion primarily when the audience is unmotivated, unengaged or unskilled. They can influence persuasion indirectly in the central, analytic route through effects on processing and metacognition. And, exceptionally, source information can be an explicit argument. Such sensitivity to sources seems intuitively reasonable, but the psychology of persuasion lacks an appropriate normative theory, the credibility of sources being determined, like the strength of arguments, in pre-tests (for further discussion, see Chapter 6).

A strikingly different approach is taken in a recent psychological theory, the Argumentative Theory of Reasoning (Mercier & Sperber, 2011, 2017). This theory places sources at the heart of reasoning, decision making, and judgement, in the sense that these cognitive abilities are seen as fundamentally geared towards interaction with others. On this view, reasoning (construed broadly to include decision making and judgement) is argumentative: it is not intended for deriving conclusions but rather for '[devising] and [evaluating] arguments intended to persuade' (Mercier & Sperber, 2011, p. 57). Mercier and Sperber acknowledge evidence that, on an individual level, much of our reasoning seems to be biased or motivated (see, e.g., Kunda, 1990), but argue that this behaviour can be argumentatively appropriate. For instance, it may seem undesirable for individual reasoners to generate evidence only in support of their belief, neglecting possible

RATIONALITY, PRAGMATICS, AND SOURCES

falsification. But, Mercier and Sperber argue, such a confirmation bias primarily affects the production, not the evaluation of arguments, and allows an efficient division of labour, both sides producing as many arguments as possible and then, ideally, resolving their dispute together. Mercier and Sperber apply their theory to a wide range of classic fallacies. They argue, for instance, that people make choices that are easy to justify, regardless of their ultimate optimality, and argue that such reason-based choice can explain the attraction effect, disjunction effect, the sunk-costs fallacy, framing effects, and preference inversion (Mercier & Sperber, 2011). They also take a somewhat optimistic view of the data on group decision making, concluding that such data show that discussion improves reasoning (though, as we will below, see qualifications are due).

The claim that argumentation is social has precedent in Argumentation Theory. Research on argumentation asks what makes a good natural-language argument, whether people are good at natural-language argumentation, and whether people's skills can be improved. Argumentation theory has taken a social turn, the focus being less on relationships between true and false propositions and more on the use of argument in reasoned dialogue (Toulmin, 1958; Walton, 2008). Since the social turn, argument strength is treated as relative to an audience (Perelman & Olbrechts-Tyteca, 1969; Toulmin, 1958; Toulmin, Rieke, & Janik, 1979). Argument strength can also vary across social contexts: for example, the same argument may be strong in a casual discussion but weak in a court of law. The content and contexts of argument can demand particular criteria for validity, criteria which are established by the participants in the argument (Toulmin, 1958; Toulmin et al., 1979).

RATIONALITY, PRAGMATICS, AND SOURCES

More recent Argumentation Theory places even greater emphasis on social factors. This emphasis takes the form of rules intended to regulate the behaviour of participants. Take, for instance, the following rules from Pragma-Dialectics:

Rule 1 (Freedom Rule): Discussants may not prevent each other from advancing standpoints or from calling standpoints into question.

Rule 2 (Burden of Proof Rule): Discussants who advance a standpoint may not refuse to defend this standpoint when requested to do so

(Eemeren, Garssen, & Meuffels, 2009, pp. 21–22).

These rules can be construed as social or ethical norms, their ethical character being clearer in the ‘fairness rules’ formulation of Christmann and colleagues (Christmann, Mischo, & Flender, 2000; Christmann, Mischo, & Groeben, 2000; Schreier, Groeben, & Christmann, 1995). Such rules have also been also put to work in explaining both qualitative (Eemeren & Houtlosser, 1999) and quantitative data (Christmann, Mischo, & Groeben, 2000; Eemeren et al., 2009; Schreier et al., 1995).

Sources figure especially strongly in a range of arguments which were classically deemed fallacious but which can, since the social turn, be considered good arguments. Historically, since sources were irrelevant to argument quality, arguments were deemed poor if they relied on appeal to features of a source (*argumentum ad verecundiam*) or attacking a source (*argumentum ad hominem*). However, properties of the source do seem to matter: intuitively relevant are its expertise, accuracy, and trustworthiness. Indeed, contemporary Argumentation Theory attempts to explain when source information is relevant and how it should be judged. Take, for instance, the following argumentation scheme and critical questions for an argument from expertise from Walton, Reed, and Macagno (2008):

Argument Scheme

RATIONALITY, PRAGMATICS, AND SOURCES

Source E is an expert in subject domain S containing proposition A.

E asserts that proposition A (in domain S) is true (false).

A may plausibly be taken to be true (false).

Critical Questions

How credible is E as an expert source?

Is E an expert in the field that A is in?

What did E assert that implies A?

Is E personally reliable as a source?

Is A consistent with what other experts assert?

Is A based on evidence?

It is questionable whether such critical questions really do provide anything like a normative basis (Hahn & Hornikx, 2016). But, as this approach shows, Argumentation Theory assumes that sources should be considered.

Related to both persuasion and argumentation is the philosophy of testimony. Testimony, in the philosophical sense, is information received from other people; the study of testimony has increasing impact on the psychology of rationality (see chapters in Zenker, 2012). Theories of testimony can provide both normative accounts and hypotheses to test. For testimony, the normative question is whether it is justifiable to revise one's beliefs in response to testimony and, if so, how (see, e.g., Coady, 1992). Testimony has proved amenable to normative Bayesian models, which will be explored in Chapter 6 (Bovens & Hartmann, 2003; Olsson, 2011; Olsson & Vallinder, 2013). These models prescribe consideration of sources whenever information is received from other people; hence, they insist on a close relationship between rationality and sources. A growing literature tests whether these

RATIONALITY, PRAGMATICS, AND SOURCES

models are descriptively accurate (e.g. Hahn, Harris, & Corner, 2009; Harris & Hahn, 2009; Harris, Hahn, Madsen, & Hsu, 2016).

Testimony overlaps with many tasks in judgment and decision making. This is perhaps most obviously the case in individual decision making by description, in which an individual makes a decision based on information provided to them by some source. Sources have made their way into the literature on individual judgement and decision making: recall the examples above in which experiments manipulated sources when exploring, for instance, the false-memory effect (Dodd & Bradshaw, 1980; Smith & Ellsworth, 1987) or ‘saying things that go without saying’ (Gruenfeld & Wyer, 1992). There is, of course, a more fundamental point: experimental participants know that an experiment is designed by an experimenter to some end; they may factor this into their interpretation of materials, and into their decisions about whether to trust the information they receive (Hilton, 1995; Schwarz, 1996). There is scarce research on actual testimonial models in this context. A notable exception is Bovens and Hartmann’s (2003) treatment of the conjunction fallacy, an analysis which shows that a sophisticated probabilistic reasoner could produce the typical response pattern (rating the conjunction higher than the conjuncts) under certain assumptions. This model has been tested by Jarvstad and Hahn (2011), who found no evidence to support the model’s assumptions. This issue will resurface in Chapter 4.

Testimony also overlaps with group judgement and decision making. Although there are testimonial models of receiving information from multiple sources (e.g. Bovens & Hartmann, 2003), these models do not seem to have had an impact on the group decision-making literature. There is considerable potential for useful dialogue between the two disciplines. A key question is how individual and

RATIONALITY, PRAGMATICS, AND SOURCES

group performance are related: under what conditions does group performance outperform individual performance (for a review, see Gigone & Hastie, 1997; Kerr & Tindale, 2004)? The main focus has been on interaction: for instance, studies have had groups choose and nominate their best member (the dictator method); or freely debate to reach consensus (the consensus method); or consider the group mean, think of reasons why it might be too high or low, and then decide on an answer (the dialectic method); or just view all judgements, and make revisions in rounds, without discussion, until consensus is achieved (the Delphi method) (eg. Snizek, 1989). The crucial measure is how the group performs relative to some baseline, often the performance of the most successful member (Kerr & Tindale, 2004).

Groups vary from the baseline by making process gains (they outperform their best member) or process losses (they are outperformed by their best member). These deviations from the baseline give insight into rationality. Take, first, process losses. There are numerous examples in the literature. When there is discussion, one party can dominate, trains of thought can be prevented from arising or, if they do arise, can be derailed (Kerr & Tindale, 2004). Other common deficits are demotivation and social loafing (Kerr & Tindale, 2004). Groups can also exchange information sub-optimally. For instance, in the 'hidden profile method', some information is distributed to all participants (shared information); other information is distributed only to some (unshared information) (Stasser & Titus, 1985, 2003). Participants tend to overweight shared information, especially under time pressure; they prefer to present and receive shared information; and they judge someone who gives shared information as more competent and credible (for discussion, see Kerr & Tindale, 2004). There are, nevertheless, cases of process gains. When groups are cohesive, share strong productivity norms, and have a shared understanding of the

RATIONALITY, PRAGMATICS, AND SOURCES

task, they show high motivation, and can perform strikingly well (Mellers et al., 2014). Such factors were exploited in the recent Good Judgment Project, in which groups who discussed forecasts freely outperformed both individuals and groups who only saw other members' judgments (see, e.g., Atanasov et al., 2015; Mellers et al., 2014; Mellers, Stone, Atanasov, et al., 2015; Mellers, Stone, Murray, et al., 2015; Tetlock, Mellers, Rohrbaugh, & Chen, 2014).

Such experiments offer rich testimonial contexts and ample opportunities for models of complex testimony with multiple sources (see Bovens & Hartmann, 2003, and for discussion Hahn, Harris, & Corner, 2016). These rewarding prospects argue for further research to explore how people interact and how they process the information they receive from sources. Such fine-grained manipulations are unfeasible in such large-scale studies as the Good Judgment Project. There is, however, suggestive evidence already. Participants seem to consider sources expert if they are loquacious, influence other group members using reason, and are confident or dominant (Kerr & Tindale, 2004; for the empirical studies, see Littlepage & Mueller, 1997; Littlepage, Schmidt, Whisler, & Frost, 1995). Much remains to be done to explore the rationality of both these reliability judgements and their subsequent use in weighting evidence.

Although such group decision making is an extremely valuable topic of research, this thesis will focus on individual judgement, reasoning, and decision making by description. This focus does not imply that this is the more important level. Rather, these individual judgements mesh better with natural-language pragmatics, which has focused, thus far, on dyadic or small-scale interactions. This focus also allows us to explore basic mechanisms which can later be extended to

group contexts. As we will see in Chapters 2-4 and 6, there is much to be done at this lower level.

1.4 Prospectus

This thesis will address the relationships among the three components. Chapters 2 to 4 explore, in particular, the relationship between pragmatics and sources, though rationality will never be far from the debate. These chapters investigate conditionals in simple testimonial contexts. The data bear directly on how participants interpret conditional premises in a wide variety of experiments. The data, as a whole, are not well accommodated by the leading theories of the conditional. Chapter 5 also investigates conditionals, but this time in the context of goal framing. This chapter focuses on the link between pragmatics and rationality. The data offer new insight into goal framing by providing evidence that participants are sensitive to the utilities of the frames. Chapter 6 focuses on the link between sources and rationality, investigating how people respond to information from partially reliable sources. The data suggest that people use the expectedness of claims to calibrate their perception of a source's reliability, and that they can treat sources as anti-reliable: as negatively correlated with the truth. Chapter 7 develops the experimental method from Chapter 6 to return to the relationship between pragmatics and sources. This chapter focuses on hedging, a phenomenon which plays a key role in a recent theory of pragmatics. The data do not support this theory

Throughout, the thesis will argue for the inseparability of rationality, pragmatics, and sources. Although each chapter will foreground two of these components, the third will remain in the background. Chapter 8 will draw together the experimental data and will further develop the themes in this introductory chapter. It will, in particular, return to the topic of argumentation and use it as a case

RATIONALITY, PRAGMATICS, AND SOURCES

study for reconceptualising the links between rationality, pragmatics, and sources.

The chapter will also discuss emerging research programs which have the potential to fundamentally reshape the approach to rationality.

2 Conditionals and Testimony

Chapter 1 introduced a triangular scheme for interpreting the relationship between rationality, pragmatics, and sources. The link between pragmatics and sources is the topic of this and the following two chapters¹⁸. Each chapter treats the conditional, understood here as the construction ‘If P, (then) Q’, which has a privileged place in philosophy, psychology and linguistics. More specifically, the chapters treat conditionals in the context of testimony, in the philosophical sense. Testimony occurs when a speaker says, asserts, or tells someone something; in epistemology, testimony is studied descriptively – studying when people accept what they are told - and normatively - when they *should* accept what they are told (Adler, 2015). Testimony is fundamentally linked with pragmatics: both involve the intentional transfer of belief from a speaker to a hearer. Considering the two phenomena together forces us to think of the connection between pragmatics and sources: particularly, between pragmatics and attributes of the speaker(s) such as their reliability. Although the connection between pragmatics and sources is the principal focus of these chapters, the issue of rationality will surface repeatedly, emphasizing the closeness of all three components.

Why study conditionals in this way? Conditionals feature prominently in the psychological of rationality, and their testimonial aspects, though crucial, are underexplored. Conditionals abound in communication, since they are a natural vehicle for expressing information under uncertainty (Hilton, Kimmelmeier, & Bonnefon, 2005) and for expressing arguments from consequences, which are central to persuasion (Bonnefon, 2016; Schellens & De Jong, 2004). Such

RATIONALITY, PRAGMATICS, AND SOURCES

conditionals are testimonial: they have a source, even if that source is only implicit.

Conditionals also arise in the psychology of decision making in two key framing effects (again, with implicit sources): risky-choice framing, in the Asian disease paradigm (Tversky & Kahneman, 1981); and goal framing, where a behaviour is advocated either using the positive consequences of compliance or the negative consequences of non-compliance (for reviews, see, e.g., Gallagher & Updegraff, 2012; Levin, Schneider, & Gaeth, 1998; Rothman & Salovey, 1997). Witness the following pair of sentences:

- (1) If you decide to get HIV tested you may feel the peace of mind that comes with knowing about your health.
- (2) If you decide not to get HIV tested you may not feel the peace of mind that comes with knowing about your health.

(Apanovitch, McCarthy, & Salovey, 2003)

Conditionals are also heavily implicated in the psychology of reasoning, where conditional inference is a central topic. Wherever such experiments manipulate sources explicitly (e.g. Stevenson & Over, 2001), or can be considered communicative contexts (see Chapter 1), conditional materials can be understood as testimony. Even though conditionals seem natural and are readily understood, as we will see below, they are also somewhat mysterious: there is little consensus on their meaning; and there is still less consensus on what we learn from the saying, asserting, or telling of a conditional – henceforth, from ‘testimonial conditionals’.

In this chapter, I will briefly survey the relevant psychological literature on the meaning of the conditional: both its semantics and its pragmatics. This literature

¹⁸ These chapters report collaborative work with Professors Ulrike Hahn, Stephan

gives a basis for the study of testimonial conditionals. I will then introduce testimony and discuss its role in generating empirical predictions. Lastly, I will introduce some experiments that are a first step towards an empirical study of the testimonial conditional. As this is the first in a series of related chapters, I will postpone some important issues to later chapters for the sake of brevity: for instance, I will discuss philosophical theories of the conditional across the chapters; and I will discuss the formal issues around learning a conditional in Chapter 4.

2.1.1 Conditionals in the psychology of reasoning

Although testimonial conditionals do not seem to have been studied independently in psychology, there are nevertheless suggestive data, especially in the psychology of reasoning. These data bear on the meaning of the conditional: its semantics and, to a lesser extent, its pragmatics. In its early days, the psychology of reasoning tended to focus on abstract tasks (Over, 2016), such as the abstract Wason selection task (Wason, 1966). It is tempting to view such tasks as pure reasoning: a sentence such as ‘If A, then B’ seems to communicate little. Even with comparatively abstract tasks, though, interpretation intervenes. In modelling the abstract Wason selection task, for instance, Oaksford and Chater (1994) make an assumption about participants’ interpretations: that the conditional in the following rule is interpreted as a conditional probability: ‘If there is an A on one side of a card, then there is a 2 on the other’. In more natural tasks, pragmatics becomes more obviously relevant (Giroto, Kimmelmeyer, Sperber, & van der Henst, 2001; Sperber, Cara, & Giroto, 1995). Naturalness is one of the key aims of the ‘New Paradigm’ in

the psychology of reasoning (Over, 2016). Naturalistic tasks can more easily be understood as communicative or testimonial contexts.

One common task in the psychology of reasoning is the truth-table task, which is often taken to offer direct evidence about the meaning of the conditional. The task recalls work in truth-conditional semantics which takes the core meaning of a sentence (at least when used to express a proposition) to be given by its truth conditions: that is, the conditions in the world that make the sentence true or false (Cann, 1993). Truth-conditional semantics uses propositional logic to model the meaning of the natural-language connectives ‘not’, ‘if’, ‘and’, and ‘or’. Propositional-logical connectives are defined with truth tables. Most relevant, for present purposes, is the truth table for the material conditional, shown in Table 2.1:

Table 2.1. Truth table of material conditional

P	Q	$P \rightarrow Q$
1	1	1
1	0	0
0	1	1
0	0	1

The material conditional is true in all cases except when P is true and Q is false: it is equivalent to $(\sim P \vee Q)$. It is truth-functional - the truth of the whole is a function of the meaning of its parts – it has had prominent defenders as a theory of the natural-language conditional (see, e.g., Bennett, 2003, for discussion; for defenders, see Grice, 1989; Jackson, 1979), and it forms the basis of the mental-models account of the conditional (Johnson-Laird & Byrne, 2002).

The truth-table task tests the truth table above against rivals, thereby offering insight into the meaning of the conditional. Common formats are (1) to present participants with the four combinations of P and Q states, and ask them to decide whether they make the conditional true or false (or, sometimes, neither); (2) to have

participants construct their own truth tables; or (3) to ask which cases prove a conditional rule (Schroyens, 2010). The major rival to the material-conditional truth table is the so-called ‘defective’ truth table (for discussion, see Over, 2016), which is shown in Table 2.2.

Table 2.2. Defective truth table

p	q	$p \rightarrow q$
1	1	1
1	0	0
0	1	Irrelevant
0	0	Irrelevant

The typical finding is that ‘P and Q’ makes the conditional true and that ‘P and not Q’ makes the conditional false; this is consistent with both truth tables (Politzer, Over, & Baratgin, 2010). But the following finding is commonly reported: participants tend to say that the ‘not P’ cases make the conditional neither true nor false, when they have that option available to them, a response which suggests that the conditional best corresponds to the defective truth table (Over, 2016; Politzer et al., 2010). Such reports may overstate the evidence. In a meta-analytic review, Schroyens (2010) showed that ‘irrelevant’ was not, in fact, the majority response for false-antecedent conditions. ‘Irrelevant’ was the modal response only for conditionals with false antecedents and false consequents where the falseness was conveyed implicitly: for instance, the rule ‘If A then 2’ was labelled ‘irrelevant’ when there were cards for ‘B’ and ‘7’. Similarly, in a recent experimental study, Skovgaard-Olsen, Kellen, Krahl, and Klauer (2017) found a poor fit between their data and any major truth table, suggesting, perhaps, complex individual differences.

Although there is controversy around the evidence for the defective truth table, the supposed evidence is interpreted in two principal ways. The first is to

argue that the conditional is best modelled with a three-valued logic (Edgington, 1995; Evans & Over, 2004; Over, 2016; Pfeifer & Kleiter, 2010). The second is to argue that conditionals are best modelled with conditional probabilities. In the truth table, the columns then represent, not the truth of the statements, but the probabilities in the special case where there is certainty. This interpretation must contend with the fact that participants are prepared to use the labels ‘true’ and ‘false’: the labels must presumably be taken to express subjective agreement or disagreement with a judgment rather than committing to a strict truth value (Over, 2016). But the probabilistic interpretation can be supported by testing other, more complex truth tables which allow for more variation in uncertainty (Over, 2016).

Another common task is the conditional-inference task. Participants typically read a conditional premise of the form ‘If p then q’ and a minor premise, and then must draw or endorse inferences. The minor premises in question could be ‘p’ for *modus ponens* (If p, q; p; therefore q), ‘not q’ for *modus tollens* (If p, q; not q; therefore not p); ‘q’ for Affirmation of the Consequent (If p, q; q; therefore p); or ‘not p’ for Denial of the Antecedent (If p, q; not p; therefore not q). The resultant data suggest that participants perform, or endorse, *modus ponens* most frequently, followed by *modus tollens*, Affirmation of the Consequent, and then Denial of the Antecedent (for discussion, see, e.g., Manktelow, 2012). These data bear on learning testimonial conditionals. Most obviously, they suggest the inferences that people will likely draw from a conditional as part of an argument; hence, they suggest the learning downstream of the conditional. But they also offer indirect evidence about the information content of the conditional itself. For, if participants understand the conditional to mean the material conditional of classical logic, and if they reason using classical logical principles, then they reason poorly. Of the inferences above,

RATIONALITY, PRAGMATICS, AND SOURCES

modus ponens and *modus tollens* are classically valid: they should be endorsed by all participants. The remaining inferences are classically fallacious: they should be rejected by all participants. However, a different account better approximates participants' reasoning. This account models the conditional as the conditional probability, and treats participants' conclusions as probabilistic inferences (e.g. Oaksford, Chater, & Larkin, 2000). Although this is only indirect evidence of the information content of the conditional itself, it is nevertheless suggestive.

More direct evidence can be found in studies using judgment tasks, in which participants assess the probability of the conditional. Such studies have asked, for instance, whether the probability of the conditional corresponds to the probability of the material conditional, $P(\sim P \vee Q)$; to the probability of the conjunction, $P(P \ \& \ Q)$; or to the conditional probability, $P(Q|P)$. Their results provide considerable support for the conditional probability (Evans, Handley, Neilens, & Over, 2007; Evans, Handley, & Over, 2003; Fugard, Pfeifer, Mayerhofer, & Kleiter, 2011; Oberauer & Wilhelm, 2003; Over, Hadjichristidis, Evans, Handley, & Sloman, 2007; Politzer, Over, & Baratgin, 2010), though also some support for the probability of the conjunction (e.g. Evans et al., 2003; Fugard et al., 2011). These data suggest that the conditional probability is closely tied to the meaning of the conditional.

A special role for probabilities can be neatly accommodated within the suppositional theory of the conditional (e.g. Adams, 1965, 1970, 1975, Edgington, 1995, 2014; Evans & Over, 2004). On this theory, when we hear a conditional *If A, B*, '[we] suppose (assume, hypothesize) that *A*, and make a hypothetical judgment about *B*, under the supposition that *A*, in light of your other beliefs' (Edgington, 2014). In other words, we judge the conditional probability $P(B|A)$. This theory takes the conditional to have, not a truth value, but a conditional probability (for a recent

discussion, see, e.g., Over, 2016). This semantics is supplemented with a pragmatic account. Since ‘it can be held that a conditional is only asserted when [the conditional probability] is high enough in context’, asserting a conditional pragmatically implies that the conditional probability is high (D. Over, personal communication¹⁹, February 2, 2017)

Despite such evidence, there is still controversy about the relationship between the probability and meaning of the conditional. Some researchers argue that basic conditionals are deterministic (e.g. Goodwin, 2014; Johnson-Laird & Byrne, 2002). Other researchers argue that, although probabilities are important, there are other requirements, such as a relevance relation between antecedent and consequent (Skovgaard-Olsen, Singmann, & Klauer, 2016b) or an inferential link (Douven, 2015; Krzyżanowska, 2015). Still others point to systematic divergences between the conditional probability and the acceptability of different types of conditionals (Douven & Verbrugge, 2010). Thus, although the literature shows a close association between conditionals and the conditional probability, it vigorously debates whether probabilities are part of the semantics²⁰ of the conditional and, if they are, whether they exhaust the semantics of the conditional. Since the empirical literature has given little attention to testimonial conditionals, it is also debatable whether changes to the conditional probability exhaust the belief change resulting from the assertion of a conditional.

¹⁹ The quotation is from David Over’s talk at the ‘Learning Conditionals’ Workshop at Ludwig Maximilian’s University, Munich, February 2nd-3rd, 2017. His lecture notes, including the quotation, are available at http://www.cas.uni-muenchen.de/veranstaltungen/archiv_veranstaltung/tagungen/ws_krzyzanowska_hartmann/index.html

²⁰ If probabilities are part of the semantics, then this will be a rather different semantics than conventional truth-conditional semantics. I thank Drs Niels Skovgaard Olsen and Karolina Krzyżanowska for discussion of this point.

2.1.2 Conditionals in Psycholinguistics

Although debate continues, the psychology of reasoning clearly suggests that when we learn a conditional from testimony, we change our estimates of the conditional probability. Further evidence is offered by psycholinguistic research. For instance, the probability of the conditional predicts reading times (Haigh & Stewart, 2011), and this seems to be due to the conditional probability, not the probability of the conjunction or of the material conditional (Haigh, Stewart, & Connell, 2012). Such psycholinguistic studies tend to use more natural materials and less intrusive dependent measures than the psychology of reasoning, which reduces the risk that the data considered thus far are attributable to demand characteristics. However, psycholinguistic studies also suggest that other probabilistic quantities are important. For instance, reading times are also predicted by the probability of the antecedent (Haigh, Stewart, & Connell, 2012), a finding which raises the possibility that people may learn more from the assertion of a conditional than just the conditional probability. When conditionals are about valued actions and outcomes, then participants also show early sensitivity to conditional utilities (Haigh, Ferguson, & Stewart, 2014) and to types of speech act, such as warnings versus tips (Stewart, Haigh, & Ferguson, 2013). This is all potential information to be learnt from the assertion of a conditional.

2.1.3 Conditionals and Pragmatics

It is tempting, but ultimately unsatisfactory, to limit one's attention to the conditional probability, since it seems to offer the basis of a theory of meaning and reasoning. Other factors – we might argue - are 'just pragmatics', and should be treated elsewhere. This argument is reminiscent of a common – and often condemned – strategy in linguistics known as the 'pragmatic waste-bin' (Mey, 2001).

RATIONALITY, PRAGMATICS, AND SOURCES

We start with a formal theory which, though elegant, cannot explain certain crucial phenomena. We then declare these phenomena to be pragmatics, often solely on the grounds that they do not fit within the theory, and take no further responsibility for explaining the phenomena. Applying this strategy more generally results in a pragmatics populated by discarded, unconnected phenomena. Specifically for the conditional, the approach is unsatisfactory for three reasons. Firstly, as we will see below, there is no well-developed pragmatic theory of the conditional to explain what is leftover. Secondly, if we are interested in the process of learning a conditional, then what is learnt in a given context will not, presumably, be limited to purely semantic information. Thirdly, pragmatics is intimately bound up with reasoning.

While the first two reasons are self-explanatory, the third requires elaboration. The essence is that pragmatics is crucial for discerning the premises and the operations on those premises. Consider, for example, the following sentence:

(3) *Syntactic Structures* is Chomsky's book.

Here, semantics alone does not yield an evaluable proposition: with the semantics alone, we can only tell that there is some relation between Chomsky and the book. A mandatory pragmatic process must specify what that relationship is: the process is mandatory because it is driven by the linguistic structure; there is a slot, here provided by the genitive 's, which needs to be filled pragmatically (Recanati, 2004). Is Chomsky the book's author, borrower, or owner? Consider, now, sentence (4):

(4) If *Syntactic Structures* is Chomsky's book, then he'll be angry that you lost it.

RATIONALITY, PRAGMATICS, AND SOURCES

Pragmatics, again, is crucial. The conditional probabilities – and, hence, any subsequent reasoning - will presumably differ depending on whether Chomsky is the author, borrower, or owner, not to mention the necessary reference assignments²¹. Sometimes, semantics delivers an evaluable proposition, but that proposition may not be the intended one. Consider sentence (5):

(5) Peter has had breakfast.

Here, there is no linguistic slot to be filled, but the proposition may optionally be specified further, for instance, by adding ‘today’ (Recanati, 2004). Consider, now, sentence (6):

(6) If Peter has had breakfast, he won’t be hungry this morning.

Again, the conditional probabilities – and, hence, any subsequent reasoning - will presumably differ depending on whether the breakfast in question was eaten today, yesterday, or twenty years ago. These pragmatic concerns will apply, too, to the conditionals themselves – indeed, to any operator. If researchers want a theory of reasoning which can predict people’s behaviour across contexts, then they will need a component which explains how the operators are interpreted and why. It is important, then, to ask how people change their beliefs when they learn a conditional from assertion. Is the change localized on the conditional probability, or do beliefs change more broadly? If belief change is broader, then there may be a profound effect on reasoning: for instance, by modifying belief in the other premises in an argument.

²¹ I am not, here, committing to conditionals expressing propositions. I will assume pragmatics is required to produce an evaluable premise, however it is formally understood.

Perhaps surprisingly, there are limited data, thus far, on the pragmatics of the conditional. Much of these data come from studies on conditionals with valued antecedents and consequents, known as utility conditionals. Such studies suggest that people draw on the utilities of antecedent and consequent and on simple mentalistic assumptions to interpret conditionals, to infer other people's intentions, and to predict other people's behaviour (Bonnefon, 2009, 2012; Bonnefon, Girotto, & Legrenzi, 2011; Bonnefon, Haigh, & Stewart, 2013; Bonnefon & Hilton, 2004; Bonnefon & Sloman, 2013; Haigh & Bonnefon, 2015). Utility conditionals will be the topic of a later chapter. It suffices, for now, to observe that utility conditionals are an important subset of conditionals, and that more remains to be done on the pragmatics of the conditional *per se*.

A more general topic is invited inferences. Geis and Zwicky (1971) suggested that conditionals tend to undergo 'conditional perfection', a pragmatic enrichment, and become bi-conditionals practically automatically (Noveck & Bonnefond, 2011). This practical automaticity has since been questioned. For instance, Evans and Over (2004) have observed that the inference 'If p then q; q; therefore p' tends to occur more with abstract conditionals, and that the inference 'If p then q; not p; therefore not q' tends to occur more with everyday conditionals. There is arguably, then, no strong connection between the two inferences (for discussion, see Bonnefon & Politzer, 2010). Moreover, as Noveck and Bonnefon (2011) observe, the idea of default pragmatic inferences has lost favour as, contrary to prominent theories (e.g. Levinson, 2000), even seemingly routine inferences have been shown to require effortful contextual processing. Conditional perfection arguably arises out of the expectation that, when we hear a conditional, we will then hear information about the antecedent being satisfied (Noveck & Bonnefond, 2011).

When we do not hear such information, we display a surprise response, and attempt to accommodate this information by pragmatic enrichment (Bonnefond et al., 2012; Noveck & Bonnefond, 2011). On this account, once this accommodation has occurred, the classically fallacious inferences can still be inhibited by people with high cognitive abilities, but this process is effortful and time-consuming (Bonnefond et al., 2012; Noveck & Bonnefond, 2011).

While it is worth bearing this pragmatic account in mind, it applies in the context of reasoning: the pragmatic inferences arise out of a conflict between the expectation of information about the antecedent and the actual appearance of a minor premise which does not satisfy this expectation. There is a more fundamental question to be asked: namely, what is learnt from the assertion of a conditional. The following three chapters will address this question.

2.1.4 Testimony

Where, then, should we look for hypotheses about what is learnt from the assertion of a conditional and why? There is a rich literature on the philosophy of testimony which treats the issue of learning from an assertion. Within this literature, formal Bayesian models have been developed to prescribe how people should revise their beliefs in response to the testimony of other people (e.g. Bovens & Hartmann, 2003). Chapter 4 will treat these models in detail. It suffices, at this point, to note the following. The models in question treat belief as subjective probability, and consider the normative impact of factors such as the number of sources, the coherence of their testimony, the relationship between the sources (i.e. whether they are dependent or independent), and the reliability of the sources (Bovens & Hartmann, 2003; Olsson, 2005).

Drawing on these models not only allows us to generate hypotheses about how beliefs change and why. It also allows us to ask questions about the rationality of people's behaviour. Indeed, this literature on testimony has already generated rich hypotheses for empirical research. For example, Bovens and Hartmann (2003) offered a reanalysis of the conjunction fallacy. People are taken to commit the conjunction fallacy when they rate a conjunction as more probable than its conjuncts. In their analysis of typical experimental task, Bovens and Hartmann (2003) considered people's prior beliefs about the conjuncts and how they judge the reliability of the information source. They showed that, as long as the individual propositions are interpreted as coming from different sources – or, perhaps, are being entertained separately at different time points – then for a range of priors, the standard response is, in fact, legitimate. Although the models' predictions do not find support in experimental data (Jarvstad & Hahn, 2011) they show the value of considering the literature on testimony. Moreover, since the literature has recently turned to the specific issue of learning a conditional from testimony (e.g. Hartmann & Rafiee Rad, 2017), the empirical data below will dovetail with cutting-edge formal work. As we will see, the data also provide useful constraints on modeling.

2.1.5 Towards an experimental method

This chapter reports the first in a series of studies which explored the learning of conditionals in simple testimonial contexts. The experiments were inspired by the literature on testimony, and were loosely based on Stevenson and Over (2001), who, uniquely in the literature, explored the impact of source reliability and conversational contexts on reasoning. For present purposes, the most relevant of their experiments is the following. They presented participants with arguments from sources of different reliabilities. One item, for instance, was the following:

RATIONALITY, PRAGMATICS, AND SOURCES

Major Premise: If Bill has typhoid he will make a good recovery.

Minor Premise: Bill has typhoid.

The relevant source, here, was either a medical professor (high expertise) or a medical student (low expertise); this manipulation was applied to either the major or minor premise. Participants indicated their certainty in the conclusions of *modus ponens* and *modus tollens* arguments. Source expertise predicted participants' certainty in the conclusions, an effect which has recently been partially replicated by Singmann, Klauer, and Beller (2016): the effect was replicated for *modus ponens* and extended to Denial of the Antecedent, but did not replicate for *modus tollens*.

The present studies were adapted from Stevenson and Over's (2001) design. Participants did not derive inferences, but rather performed a judgment task. There were two versions of this task, performed by different participants; both manipulated variables from the literature on social epistemology. The first task manipulated assertion: a conditional could be asserted by multiple sources or a single source, or there was a null condition. For instance:

Multiple Condition Adam, Barbara, Nick and Sue are at a large car dealership. They tell you, 'If a car is a Mercedes, then it's black.'

Single Condition Adam is at a large car dealership. He tells you, 'If a car is a Mercedes, then it's black'

Null Condition Imagine you are at a large dealership.

In each case, participants were asked about the probability that a car on the lot was a Mercedes; that a car on the lot was black; and that the car on the lot was black given that it was a Mercedes. The design was between-subjects, but each participant provided estimates of all three probabilities for each item. The conditionals

RATIONALITY, PRAGMATICS, AND SOURCES

themselves were simple indicative conditionals with neutral content²². They were designed to avoid obvious real-world causal relations, to avoid noise from participants' different causal beliefs and knowledge.

A different group of participants performed a second task, which manipulated source expertise: a conditional could be asserted by an expert or inexpert source. All but one of the conditionals from the first task were used, with different contexts. For instance:

Expert Condition Imagine you are at an infectious-diseases ward. A professor of medicine tells you, 'If Bill has malaria, then he'll make a good recovery'.

Inexpert Condition Imagine you are at an infectious-diseases ward. A medical student tells you, 'If Bill has malaria, then he'll make a good recovery.'

The same probability questions were asked as above.

The purpose of this study was exploratory: we did not have specific predictions about how the probability of the antecedent and consequent would change. A subsidiary aim was to interpret the data in the light of philosophical theories of the conditional. For expository reasons, these theories will largely be discussed in later chapters. The predictions, then, are non-directional.

Of the two tasks, the assertion task is more obviously pragmatic. In particular, the single-assertion condition tests the pure case of learning a conditional from an assertion. The source-expertise task is more obviously testimonial, in that it manipulates a variable which is theoretically important for the literature on

²² Such conditionals might be termed 'contingent universals' (Douven, 2015, p. 16),

testimony but which has played little role in the pragmatics literature (though see McCready, 2014; Sperber et al., 2010). However, I will not attempt to firmly distinguish pragmatics and testimony here; the extent to which this is possible is a question I will return to at various points throughout the thesis.

2.2 Experiment 2.1: effect of assertion

2.2.1 Methods

Participants. 179 participants (68 female; mean age 32.81) completed this experiment; 5 participants had already been excluded since they were non-native English speakers. Participants were recruited via the intermediary MTurk Data (www.mturkdata.com). High qualifications were set for the task to improve the quality of the data and maximize the number of native English speakers: participants had to be resident in the US, and have an overall approval rating of 99% for 1,000 previously completed tasks. Participants received a small fee, chosen to exceed the US minimum wage per minute.

Materials. Participants took web surveys on Amazon Mechanical Turk. The items were as follows. Note that the items are presented in the multiple-assertion condition, with the single-assertion condition in brackets. In the null condition, the text read, instead, 'Imagine you are at a large dealership (or infectious diseases word etc.), and the conditional was not presented.

(1) Adam, Barbara, Nick and Sue (Adam) are at a large car dealership.

They say, 'If a car on this lot is a Mercedes, then it's black.'

as they summarize real-world regularities without expressing temporal or causal information.

RATIONALITY, PRAGMATICS, AND SOURCES

- (2) Olivia, Aaron, Zoe and Felix (Olivia) are at an infectious diseases ward. They say, 'If a patient on this ward has malaria, then he'll make a good recovery.'
- (3) Jack, Isabel, Elliot and Amy (Jack) are at a university halls of residence. They say, 'If Lisa, a student, has an essay to write, then she is studying late in the library.'
- (4) Sophie, Owen, Julia and Nathan (Sophie) are at a veterinary hospital. They say, 'If Fido the dog is in the hospital kennels, then he's had his operation.'
- (5) Harry, Emma, Katie and Rory (Harry) are at a butterfly house. They say, 'If a butterfly is purple, then it's a fast flyer.'
- (6) Lily, Tom, Sarah and Leo (Lily) are at a New York coffee shop. They say, 'If Dan, a customer, is drinking coffee, then his drink is decaffeinated.'
- (7) Josh, Holly, Jamie, Maria (Josh) are at a restaurant. They say, 'If the dish is chicken stew, then it's gluten free.'

These items were followed by questions about the probability of the antecedent, the probability of the consequent, and the conditional probability. In the null condition, the same events were used, but there was no conditional asserted. The probability questions were simply 'What's the probability that [antecedent/consequent/consequent given antecedent²³]?'

²³ That is, the actual events were used, not the words 'antecedent', 'consequent' and so on.

RATIONALITY, PRAGMATICS, AND SOURCES

Procedure. The design manipulated Assertion (Null, Single Assertion, Multiple Assertion); the dependent measures were the probability of the antecedent, $P(\text{Antecedent})$; the probability of the consequent, $P(\text{Consequent})$; and the conditional probability, $P(\text{Consequent}|\text{Antecedent})$. The task was between-subjects: participants were assigned, in a round-robin fashion, to a condition. After giving informed consent, participants gave demographic information and were then shown the following instructions:

‘Thank you for taking part in this study. You will read about some people who are making claims. You will be told who is making the claim and what they are saying. For example, you might read ‘If the ferry is painted blue, then it’s called The Empress.’ Below each claim you’ll be asked a question. You will be asked to judge a probability on a scale from 0 (not at all possible) to 10 (certain). But each question is subtly different, so please read the wording carefully. There are no correct answers. Please just give the probability that comes to mind.’

The conditional above was included for illustration; it was not an experimental item. Participants then provided ratings for each probability question for each item on a Likert-style scale from 0 (not at all possible) to 10 (certain). There was a different page for each combination of the conditional and probability question: that is, participants saw each item three times, with a different probability question. The order of the pages was randomized. Finally, participants received debriefing information.

2.2.2 Results & Discussion

For all data in this chapter, the analyses comprised crossed random-effects models. The analyses treated the probability questions as separate dependent variables. These analyses followed the guidelines of Barr, Levy, Scheepers, and Tily (2013). Significance was determined by a likelihood ratio test (maximum-likelihood method) on two models: the full model, which included the fixed effect in question; and the null model. Unless otherwise specified, both models included the full random-effects structure justified by the design. This structure included the random slopes of the independent variable across items, random intercepts of items, and random intercepts of participants. The covariance of random slopes and items was also estimated. It is perhaps more common to test for the significance of random effects by, for instance, starting with a full model and simplifying it by removing non-significant effects (Finch, Bolin, & Kelley, 2016; Tabachnick & Fidell, 2013). However, as Barr, Levy, Scheepers and Tily (2013) note, experimental designs rarely have enough power to be confident in rejecting the random effects, and maintaining the full random-effects structure, as far as convergence allows, yields models that have been shown to generalize better in Monte Carlo simulations and which have a low risk of Type 1 errors.

Alongside the likelihood-ratio tests are estimates of the (unstandardized) coefficients for the fixed effects and the 95% confidence intervals (calculated using the Wald method) for these estimates. These coefficients for fixed effects serve as estimates of effect size²⁴. Since the aim, throughout, is to control for random effects, coefficients for the random effects are not relevant for present purposes and will not

be reported. Analyses were run using the R programming language (R Core Team, 2016) and the lme4 package (Bates, Maechler, Bolker, & Walker, 2015); where relevant, pairwise comparisons were run using the lsmeans package (Leith, 2016).

Probability of the Antecedent. Figure 2.1 shows the descriptive data for the effect of Assertion on the probability of the antecedent²⁵.

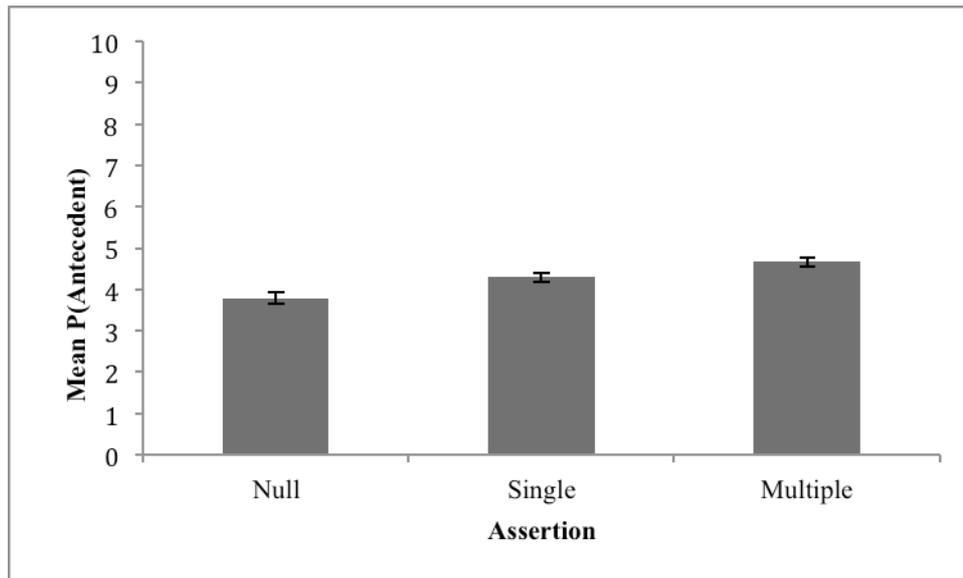


Figure 2.1. Mean rating of P(Antecedent) by condition; error bars are standard error

Note that, although the data suggest an increase from null ($M = 3.79$) to single ($M = 4.29$) and multiple ($M = 4.67$) conditions, the difference between means is small.

Indeed, including Assertion did not significantly improve fit²⁶ over the null model, $\chi^2(2) = 4.55, p = .10$. Table 2.3 summarizes the estimates of the fixed effects.

²⁴ Unit-free measures, such as r , are not straightforwardly calculable with crossed random-effects models and can produce uninterpretable results (Tabachnik & Fidell, 2013).

²⁵ For all reported crossed random-effects models, I present figures which show the descriptive data by trial; the figures do not take account of nesting. However, if we compare these descriptives with the unstandardized coefficients and least-squares means, we can see that the figures provide a not unreasonable summary. For present purposes, the chosen figures are more illustrative than the more complex plots discussed, for instance, by Finch, Bolin, and Kelley (2016).

²⁶ Note that here, and throughout, assumptions were met, unless otherwise specified.

Table 2.3. Fixed effects of Assertion on P(Antecedent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 3.79$	2.84, 4.74
Single	$b = .50$	-.24, 1.24
Multiple	$b = .88$	0.13, 1.62

Probability of the Consequent. Figure 2.2 shows the descriptive data for the effect of Assertion on the probability of the consequent.

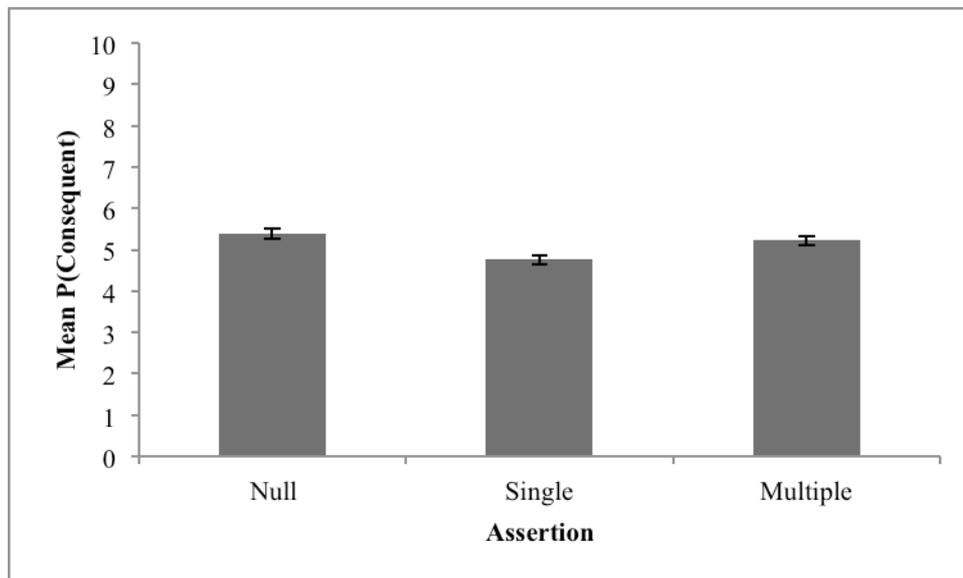


Figure 2.2. Mean rating of P(Consequent) by condition; error bars are standard error

These data do not suggest an obvious trend. The mean rating decreases from the null ($M = 5.39$) to the single conditions ($M = 4.77$) and then increases to the multiple condition ($M = 5.23$). As above, mean ratings are close together. Also as above, including Assertion did not significantly improve fit over the null model, $\chi^2(2) = 5.13, p = .08$. Table 2.4 summarizes the estimates of the fixed effects.

RATIONALITY, PRAGMATICS, AND SOURCES

Table 2.4. Fixed effects of Assertion on P(Consequent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 5.39$	4.32, 6.46
Single	$b = -.62$	-1.44, .20
Multiple	$b = -.16$	-1.06, .73

The confidence intervals for parameter estimates for single and multiple assertion both comfortably include zero, thus offering no clear evidence for an effect.

Conditional Probability. Figure 2.3 shows the descriptive data for the effect of Assertion on the conditional probability, $P(\text{Consequent}|\text{Antecedent})$.

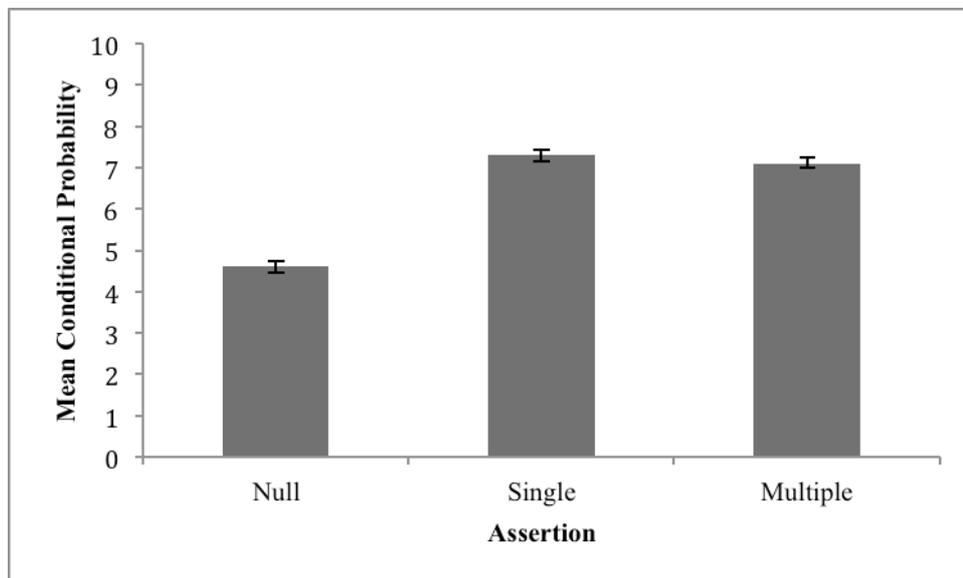


Figure 2.3. Mean rating of Conditional Probability by condition; error bars of standard error

These data suggest a trend for Assertion to increase ratings of the conditional probability. There is an increase from the null ($M = 4.60$) to single condition ($M = 7.29$), although there is then a small decrease to the multiple condition ($M = 7.10$).

This time, including Assertion significantly improved fit over the null model, $\chi^2(2) = 11.55, p < .001$. Table 2.5

Table 2.5 summarizes the estimates of the fixed effects.

Table 2.5. Fixed effect of Assertion on Conditional Probability

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 4.61$	3.54, 5.67
Single	$b = 2.69$	1.32, 4.06
Multiple	$b = 2.50$	1.43, 3.56

Pairwise comparisons were conducted on least-square means with a Tukey

correction for multiple comparisons²⁷. The increase from null to single conditions was significant, $M = 2.69$, 95% CI [.68, 4.69], $t(11.04) = 3.63$, $p = .01$, though note the wide confidence intervals, hence the uncertainty, around the estimated mean. The increase from null to multiple conditions was also significant, $M = 2.50$, 95% CI [1.02, 3.97], $t(15.87) = 4.38$, $p = .001$. The difference between single and multiple conditions was non-significant, $M = -.19$, 95% CI [-1.12, .73], $t(50.52) = -.50$, $p = .87$. This suggests that the effect of Assertion plateaus.

These data suggest that Assertion targets the conditional probability, leaving the probability of the antecedent and consequent unaffected. The data also suggest that Assertion has a diminishing return on the conditional probability: assertion by multiple sources does not increase the conditional probability. Since these are novel data, however, they require replication. This replication is reported next.

2.3 Experiment 2.2: replicating the effect of Assertion

2.3.1 Method

Participants. 188 (88 female; 1 non-binary; average age 39.02 years) completed this experiment; 1 participant had already been excluded since they were

a non-native speaker of English. Participants were, again, recruited on Amazon Mechanical Turk using the same criteria and remuneration as for Experiment 2.1.

Materials. The items were largely the same as those used for Experiment 2.1. Some small adjustments were made to one item. The conditional ‘If Dan is drinking a coffee, then his drink is decaffeinated’ creates an ambiguity. In the conditions in which a conditional is asserted, ‘his drink’ may be interpreted as the aforementioned coffee: thus, when asked about the probability of the consequent (that Dan’s drink is a coffee), participants may judge the probability of a decaffeinated (quite low). However, in the null condition, ‘his drink’ can be interpreted either as a coffee, when judging the conditional probability, or as *any* drink when judging the probability of the consequent. The probability of a drink being decaffeinated is rather different from the probability of a coffee being decaffeinated. Although the analysis controls for the effects of items, the replication removed this ambiguity by a re-wording: ‘If Dan is drinking a coffee, then it’s decaffeinated’. The P(Consequent) question referred to the probability of a coffee being decaffeinated, and the conditional probability question was phrased as follows, ‘Given that Dan’s drink is a caffeinated, what’s the probability that it is a coffee?’ The items were otherwise the same.

Procedure. The same procedure was used as for Experiment 2.1 except for the following change. The repeated occurrence of the same items may have caused confusion. To improve the presentation, all probability questions were presented

²⁷ For clarity’s sake, I report the differences between means although they largely duplicate the information from the coefficients. More important are the confidence intervals,

together for the same item. The order of the questions was reversed for half the participants. As above, participants provided ratings for each probability question for each item on a Likert-style scale from 0 (not at all possible) to 10 (certain)

2.3.2 Results & Discussion

Probability of the Antecedent. Figure 2.4 shows the descriptive data for the effect of Assertion on the probability of the antecedent.

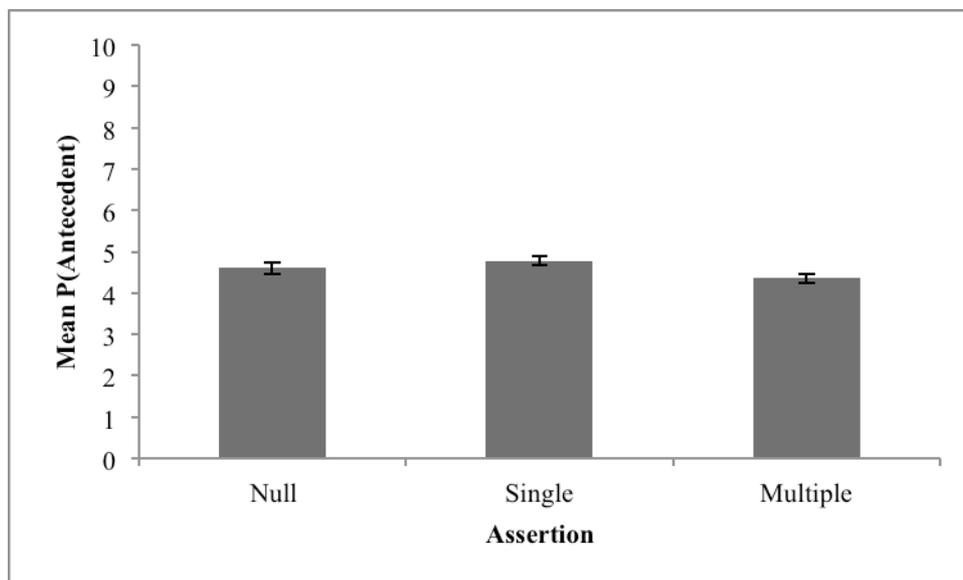


Figure 2.4. Mean P(Antecedent) by condition; error bars are standard error

These data do not suggest an obvious trend. There is a small increase in rating from the null condition ($M = 4.60$) to the single condition ($M = 4.77$), and a small decrease to the multiple condition ($M = 4.36$). Including Assertion did not significantly improve fit over the null model, $\chi^2(2) = 3.03$, $p = .22$. Table 2.6 reports the estimates of the fixed effects.

which differ from those of the coefficients because of computational differences such as a multiplicity correction.

RATIONALITY, PRAGMATICS, AND SOURCES

Table 2.6. Fixed effect of Assertion on P(Antecedent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 4.60$	3.12, 6.09
Single	$b = .16$	-.10, 1.33
Multiple	$b = -.24$	-1.26, .77

In terms of overall significance, the replication is consistent with the original data set. The parameter estimates for the single assertion condition are also consistent. The estimate for the original study was larger but still small, $b = .50$. There is also considerable overlap (1.34) between the replication confidence interval (see above) and that of the original study, 95% CI [-0.24, 1.24], and both confidence intervals include 0. The estimate for the multiple assertion condition was also larger in the original study, $b = .88$, but, again, both estimates are small. There is also overlap in the confidence intervals for the replication (see above) and original study, 95% CI [.13, 1.62], and both intervals include 0. The overlap (.64) is somewhat smaller than that for the single assertion condition, but is nevertheless large enough to consider the data consistent, at over half the length of the average margin of error of .88.²⁸ (Cumming, 2012, p. 158).

Probability of the Consequent. Figure 2.5 shows the descriptive data for the effect of Assertion on the probability of the consequent.

²⁸ As Cumming (2012) notes, if the confidence intervals overlap by up to approximately half the length of the average margin of error, then the data are significantly different at the $p < .05$ level.

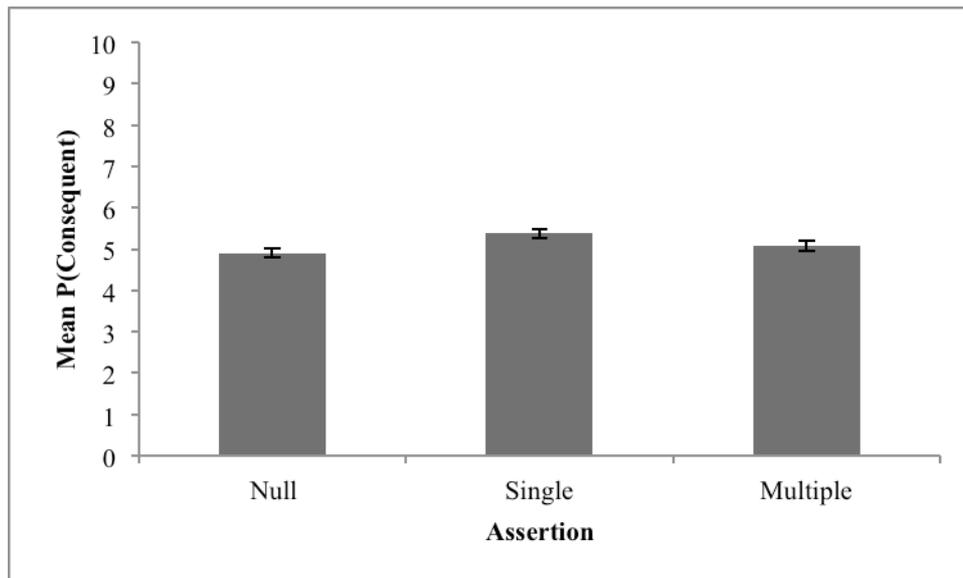


Figure 2.5. Mean P(Consequent) by condition; error bars are standard error

Once again, there is little suggestion of an effect, the means for the null ($M = 4.90$), single ($M = 5.38$), and multiple ($M = 5.08$) conditions being close together. Including Assertion did not significantly improve fit over the null model, $\chi^2(2) = 2.41, p = .30$.

Table 2.7 reports the estimates of the fixed effects.

Table 2.7. Fixed effect of Assertion on P(Consequent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 4.90$	3.98, 5.82
Single	$b = .47$	-.74, 1.69
Multiple	$b = .18$	-1.21, 1.57

In terms of overall significance, the original and replication studies are broadly consistent. The parameter estimates for the single condition are also consistent.

While the original study produced a parameter estimate with a negative sign, $b = -.62$, the confidence intervals for the original study - 95% CI [-1.44, 2] - overlap with those for the replication (see above) by some .76, which is above a benchmark of approximately half the average margin of error 1.02. The replication confidence interval is somewhat broader, suggesting greater uncertainty in the estimate. The parameter estimates for the multiple assertion condition are also consistent. Again,

the original study produced an estimate with a different sign, $b = -.16$. But the confidence intervals for the original study – 95% CI [-1.06, .73] – are fully within those for the replication. As above, the greater width of the interval does, however, suggest considerably more uncertainty in the estimate for the replication study than for original study.

Conditional Probability. Figure 2.6 shows the descriptive data for the effect of Assertion on the conditional probability.

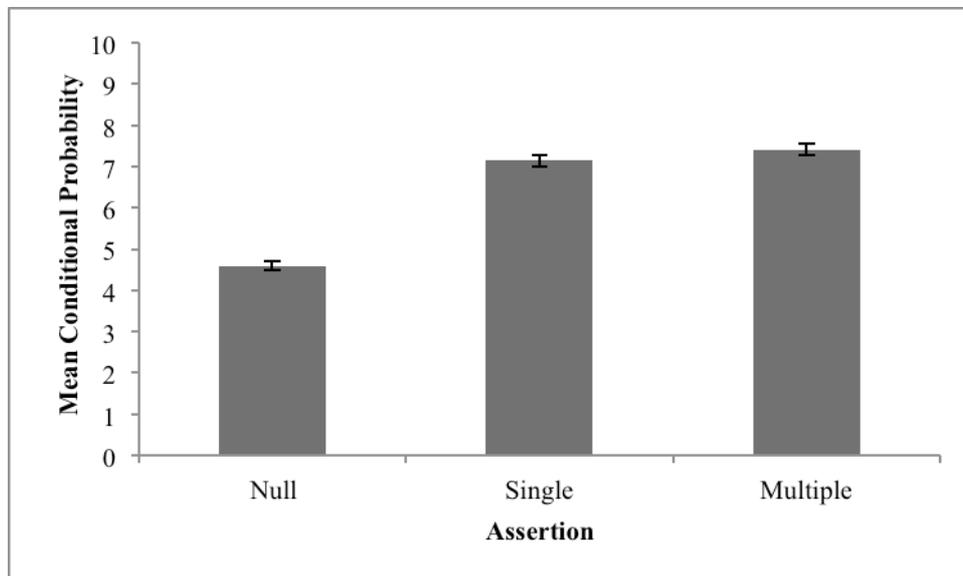


Figure 2.6. Mean Conditional Probability by condition; error bars are standard error. The data suggest a linear trend of Assertion. Mean ratings increase from the null condition ($M = 4.60$) to the single condition ($M = 7.14$) and the multiple condition ($M = 7.41$), although the difference is small between the last two conditions. Including Assertion significantly improves fit over the null model, $\chi^2(2) = 17.41, p < .001$, replicating the overall significance of the original study. Table 2.8 summarizes the estimates of fixed effects.

RATIONALITY, PRAGMATICS, AND SOURCES

Table 2.8. Fixed effects of Assertion on Conditional Probability

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 4.60$	3.75, 5.45
Single	$b = 2.54$	1.75, 3.34
Multiple	$b = 2.82$	1.90, 3.74

The original and replication studies produced consistent estimates of the fixed effect of single assertion. The original study showed a similar parameter estimate, $b = 2.69$, and the confidence interval for the replication study falls wholly within that for the original study, 95% CI [1.32, 4.06]. This suggests greater certainty about the estimate for the replication study. The estimates for multiple assertion are also consistent. The original study showed a similar parameter estimate, $b = 2.50$, and there is considerable overlap (1.66) between the confidence intervals for the replication study and those for the original study, 95% CI [1.43, 3.56].

Pairwise comparisons were performed on the least-square means, using the Tukey correction for multiple comparisons. The increase from the null to single conditions was significant, $M = 2.54$, 95% CI [1.53, 3.55], $t(46.55) = 6.09$, $p < .001$. This replicates the comparison from the original study, which showed a similar mean, $M = 2.69$ and overlapping confidence intervals, 95% CI [.68, 4.69]. The confidence interval for the replication study falls wholly within that for the original study, suggesting greater certainty in the estimate for the replication study. The increase from null to multiple conditions was also significant, $M = 2.82$, 95% CI [1.60, 4.03], $t(24.98) = 5.78$, $p < .001$. This, too, replicates the comparison from the original study, which showed a similar mean, $M = 2.50$, and overlapping confidence intervals, 95% HDI [1.02, 3.97], the overlap being 2.37. The increase from single to multiple conditions was not significant, $M = .27$, 95% CI [-.57, 1.12], $t(103.18) = .77$, $p = .72$.

RATIONALITY, PRAGMATICS, AND SOURCES

This replicates the comparison from the original study. Here, though, the sign of the mean was different, $M = -.19$. However, the original confidence intervals, 95% CI [-1.12, .73], show considerable overlap (1.3). Both sets of confidence intervals include 0. Both sets of data suggest that the effect of Assertion plateaus.

The original and replication studies offer consistent data. Although there were some differences in the means for $P(\text{Antecedent})$ and $P(\text{Consequent})$, there was substantial overlap in the confidence intervals. The data for the conditional probability showed similar means and, again, substantially overlapping confidence intervals. Together, the studies suggest that the effect of Assertion is focused on the conditional probability. The analyses for the two studies produce consistent results. There is no reliable evidence for an effect on $P(\text{Antecedent})$ and $P(\text{Consequent})$. Before discussing the implications of these data, we turn to the source expertise task.

2.4 Experiment 2.3: effect of source expertise

2.4.1 Method

Participants. 122 (50 female; average age 37.15 years) completed this experiment; 4 participants had already been excluded since they were non-native speakers of English. Participants were recruited on Amazon Mechanical Turk using the same criteria and remuneration system as for Experiments 2.1 and 2.2.

Materials. The items were as follows. Note that the items are presented in the expert source condition, with the inexpert source condition in brackets. The conditionals were the same as for Experiment 2.1, excluding item (3): during pre-testing of the items, this item did not work well under a source expertise manipulation. The items, then, were as follows:

RATIONALITY, PRAGMATICS, AND SOURCES

- (1) Imagine you are at a large car dealership. The manager (A customer) tells you, 'If a car on this lot is a Mercedes, then it's black.'
- (2) Imagine you are visiting an infectious diseases ward. A professor of medicine (a medical student) tells you, 'If a patient on this ward has malaria, then they'll make a good recovery.'
- (3) Imagine you are at a veterinary hospital. A veterinary nurse (delivery man) tells you, 'If Fido the dog is in the hospital kennels, then he's had his operation.'
- (4) Imagine you are at a butterfly house. A butterfly keeper (five-year-old visitor to the butterfly house) tells you, 'If the butterfly is purple, then it's a fast flyer.'
- (5) Imagine you are at a New York coffee shop. The girlfriend of Dan, a customer, (an old friend of Dan, another customer, who hadn't seen Dan for 10 years) tells you, 'If Dan is drinking a coffee, then his drink is decaffeinated.'
- (6) Imagine you are at a restaurant. The head chef of the restaurant (the restaurant's regular handyman) tells you, 'If a dish is chicken stew, then it's gluten free'.

As above, the items were followed by questions to elicit people's judgments about the probability of the antecedent, the probability of the consequent, and the conditional probability. The probability questions were, again, simply, 'What's the probability that [antecedent/consequent/consequent given antecedent]?'

Procedure. The design manipulated Expertise (Inexpert Source, Expert Source); the dependent measures were the probability of the antecedent,

RATIONALITY, PRAGMATICS, AND SOURCES

$P(\text{Antecedent})$; the probability of the consequent, $P(\text{Consequent})$; and the conditional probability, $P(\text{Consequent}|\text{Antecedent})$. The task was between-subjects: participants were assigned, in a round-robin fashion, to a condition. After giving informed consent and demographic information, participants were shown the same instructions as for Experiment 2.1. Participants then provided ratings for each probability question for each item on a Likert-style scale from 0 (not at all possible) to 10 (certain). There was a different page for each combination of the conditional and probability question: that is, participants saw each item three times, with a different probability question. The order of the pages was randomized. Finally, participants received debriefing information.

2.4.2 Results & Discussion

Probability of the Antecedent. Figure 2.7 shows the descriptive statistics for the effect of Expertise on the probability of the antecedent.

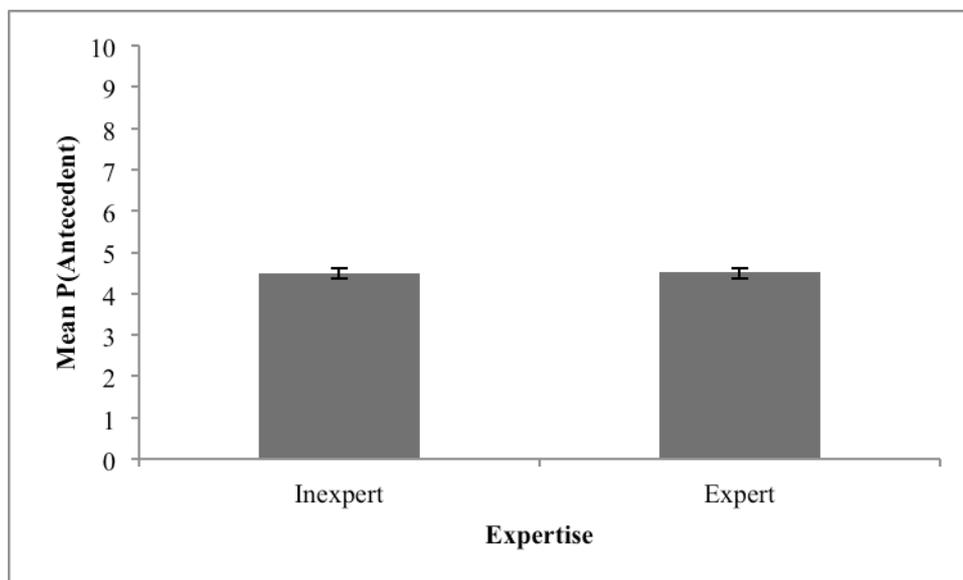


Figure 2.7. Mean $P(\text{Antecedent})$ by condition; error bars are standard error

These data do not support an effect of Expertise: note that means for inexpert sources ($M = 4.48$) and expert sources ($M = 4.51$) are very close. Indeed, including

RATIONALITY, PRAGMATICS, AND SOURCES

Expertise did not significantly improve fit over the null model, $\chi^2(1) = .01, p = .93$.

Table 2.9 summarizes the estimates of the fixed effects.

Table 2.9. Fixed effects of Expertise on P(Antecedent)

Fixed Effect	Parameter	95% CI
Intercept (Inexpert)	$b = 4.48$	3.85, 5.12
Expert	$b = .02$	-.52, .57

Probability of the Consequent. Figure 2.8 shows the descriptive statistics for the effect of Expertise on the probability of the consequent.

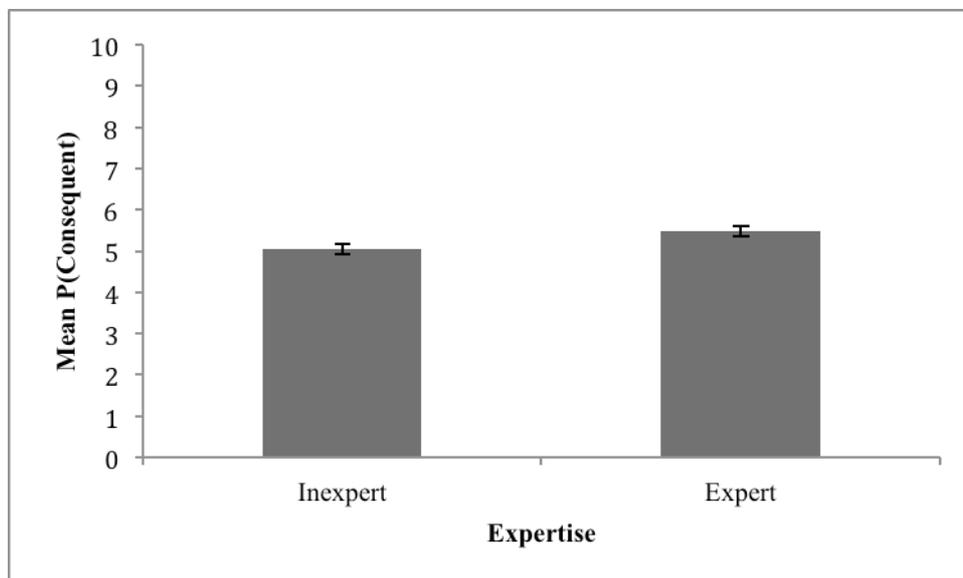


Figure 2.8. Mean P(Consequent) by condition; error bars are standard error

The means for inexpert ($M = 5.03$) and expert ($M = 5.49$) conditions are close together. Including Expertise did not significantly improve fit over the null model, $\chi^2(1) = 3.28, p = .07$. Table 2.10 reports the estimates of the fixed effects.

Table 2.10. Fixed effects of Expertise on P(Consequent)

Fixed Effect	Parameter	95% CI
Intercept (Inexpert)	$b = 5.03$	4.31, 5.76
Expert	$b = .46$	-.03, .94

Note that the lower bound of the 95% confidence interval shows only very slight overlap with 0, arguing for caution. Nevertheless, the difference between means, at .46, is small.

Conditional Probability. Figure 2.9 shows the descriptive data for the effect of Expertise on the conditional probability.

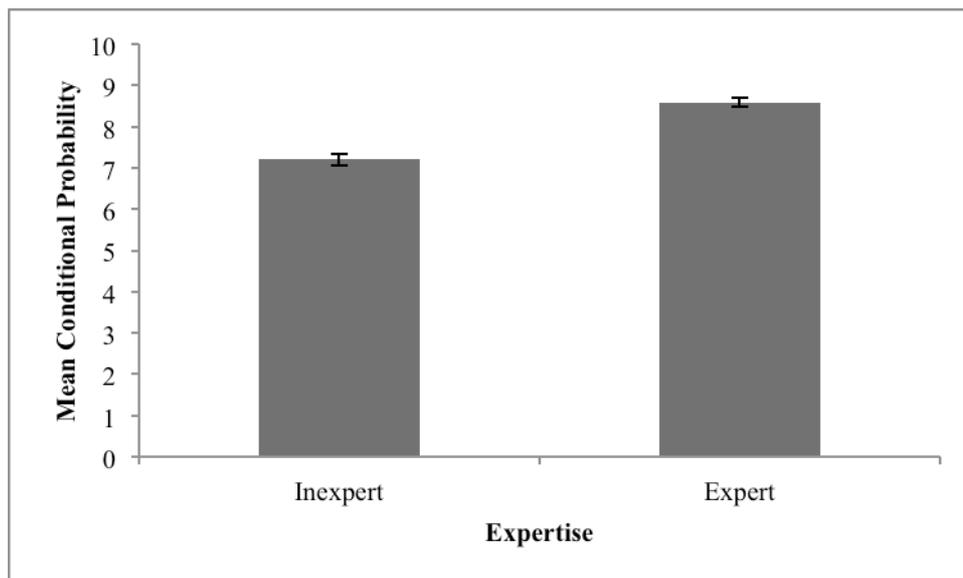


Figure 2.9. Mean Conditional Probability by condition; error bars are standard error

Mean rating of Conditional Probability by condition; error bars are standard error

The data suggest an effect of expertise: there is an increase from inexpert sources ($M = 7.23$) to expert sources ($M = 8.59$). Including Expertise significantly improved fit over the null model, $\chi^2(1) = 8.18, p < .01$. Table 2.11 reports the estimates of the fixed effects.

Table 2.11. Fixed effects of Expertise on Conditional Probability

Fixed Effect	Parameter	95% CI
Intercept (Inexpert)	$b = 7.21$	6.58, 7.84
Expert	$b = 1.38$.60, 2.15

The data for this experiment suggest that Expertise, like Assertion, has a focused effect on the conditional probability; there was no evidence of a reliable effect on the probability of the antecedent or the probability of the consequent. The novelty of these data argues for replication. Moreover, since Experiment 2.3 did not include a null condition, the data are not directly comparable with the data for Experiments 2.1 and 2.2. This defect was remedied in Experiment 2.4.

2.5 Experiment 2.4: replicating the effect of source expertise

2.5.1 Method

Participants. 179 participants (75 female; average age 36.32 years) completed this experiment; 1 participant had already been excluded since they were a non-native speaker of English. Participants were recruited on Amazon Mechanical Turk using the same criteria and remuneration system as for Experiments 2.1, 2.2 and 2.3.

Materials. The same items were used as for Experiment 2.3 with one small change. As in Experiment 2.2, the conditional (5) was changed to ‘If Dan is drinking a coffee, then it’s decaffeinated’. The P(Consequent) question referred to the probability of a coffee being decaffeinated, and the conditional probability question was phrased as follows, ‘Given that Dan’s drink is a caffeinated, what’s the probability that it is a coffee?’ A null condition was also introduced which matched that of Experiments 2.1 and 2.2: participants in this condition saw the context, such as ‘Imagine you are at a large car dealership’, and then the probability questions without seeing the conditional.

Procedure. As with the previous experiments, the survey was hosted on Mechanical Turk. The experiment manipulated Expertise (Null, Inexpert Source, Expert Source). Participants were assigned to condition round-robin style. The same procedural changes were introduced as for Experiment 2.2: there was now one page for each item. Below the item, the probability questions appeared together; the order was reversed in half the surveys. Participants rated the probabilities on a Likert-style scale from 0 (not at all possible) to 10 (certain).

2.5.2 Results & Discussion

Probability of the Antecedent. Figure 2.10 shows the descriptive data for the effect of Expertise on the probability of the antecedent.

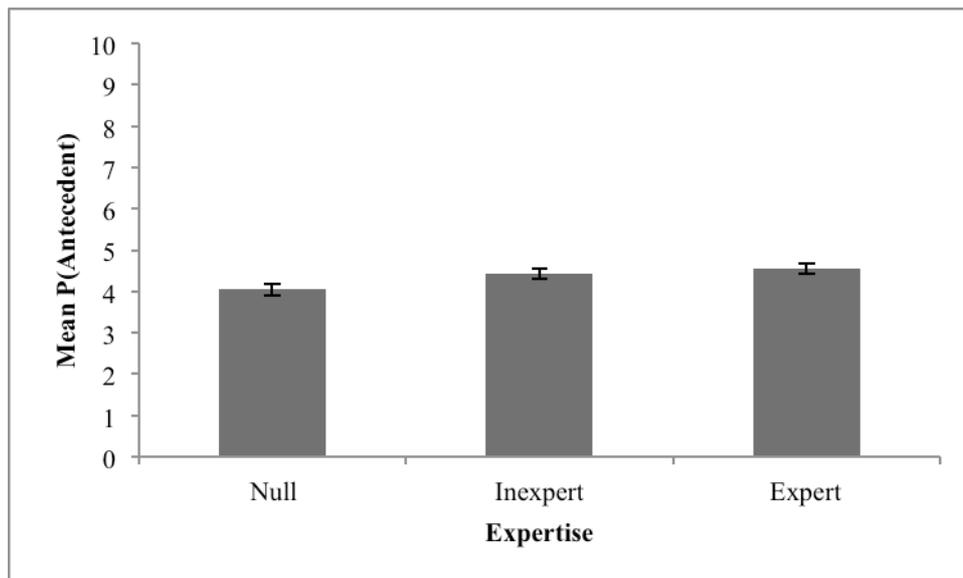


Figure 2.10. Mean P(Antecedent) by condition; error bars are standard error

The data suggest little evidence for a clear effect, although there is a very small increase from the null condition ($M = 4.05$) to the inexpert condition ($M = 4.43$) and the expert condition ($M = 4.56$). Indeed, including Expertise did not significantly improve model fit, $\chi^2(2) = .77, p = .68$. Table 2.12 reports the estimates of the fixed effects.

Table 2.12. Fixed effects of Expertise on P(Antecedent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 4.05$	2.58, 5.51
Inexpert	$b = .38$	-.78, 1.54
Expert	$b = .52$	-.67, 1.71

In terms of overall significance, the data replicate the original study. In the original study, the Expertise variable had only two levels, whereas here there are three. However, parameter estimates can still be compared by reconstructing the relevant values, since the b values correspond to the mean in the baseline condition, for the intercept, and the mean difference from it for the remaining conditions. The intercept in the original is the estimate for the inexpert condition, $M_{Inexpert} = 4.48$, 95% CI [3.84, 5.12]. This corresponds to the replication inexpert condition (reconstructed from the table above), $M_{Inexpert} = 4.43$, 95% CI [3.27, 5.59]. The parameter estimates are tightly clustered, and the confidence interval for the original inexpert condition is wholly included in the confidence interval for the replication inexpert condition. The estimates are, thus, consistent. The estimates for the expert condition are closely clustered, and the confidence interval for the replication once again wholly include that for the original study: for the original study, $M_{Expert} = 4.50$, 95% CI [3.96, 5.05]; for the replication, $M_{Expert} = 4.57$, 95% CI [3.38, 5.76].

Probability of the Consequent. Figure 2.11 shows the descriptive data for the effect of Expertise on the probability of the consequent.

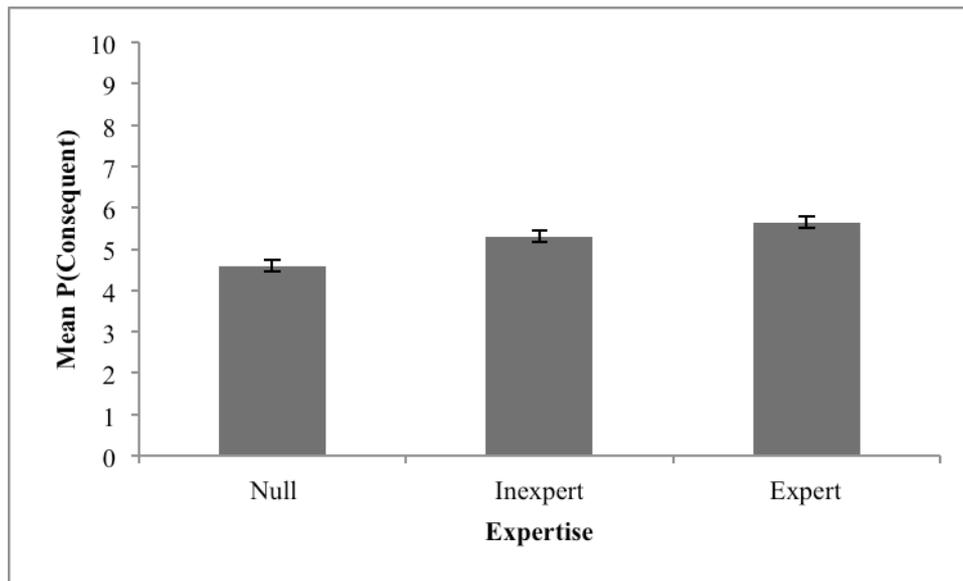


Figure 2.11. Mean P(Consequent) by condition; error bars are standard error

The data suggest a small increase from the null condition ($M = 4.60$) to the inexpert condition ($M = 5.30$) and the expert condition ($M = 5.65$). However, including Expertise did not significantly improve fit over the null model, $\chi^2(2) = 1.34, p = .51$.

Table 2.13 reports the estimates of the fixed effects.

Table 2.13. Fixed effects of Expertise on P(Consequent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 4.60$	3.54, 5.66
Inexpert	$b = .70$	-.88, 2.29
Expert	$b = 1.05$	-.84, 2.93

The replication study produced a non-significant result, as did the original study. The estimate for the fixed effect of the inexpert condition in the original study was $M_{Inexpert} = 5.03, 95\% \text{ CI } [4.31, 5.76]$. The estimate for the inexpert condition above is $M_{Inexpert} = 5.30, 95\% \text{ CI } [3.72, 6.89]$. The estimates are close together, and the confidence interval for the replication wholly includes that for the original study, suggesting consistency. The estimates for the expert condition are very close for the original and replication studies, and, once again, the confidence interval for the replication study wholly includes that for the original study: for the original study,

$M_{Expert} = 5.49$, 95% CI [5.00, 5.97]; and for the replication, $M_{Expert} = 5.65$, 95% CI [3.76, 7.53]. Note that the latter interval is rather wide, suggesting uncertainty about the estimate.

Conditional Probability. Figure 2.12 shows the descriptive data for the effect of Expertise on the conditional probability.

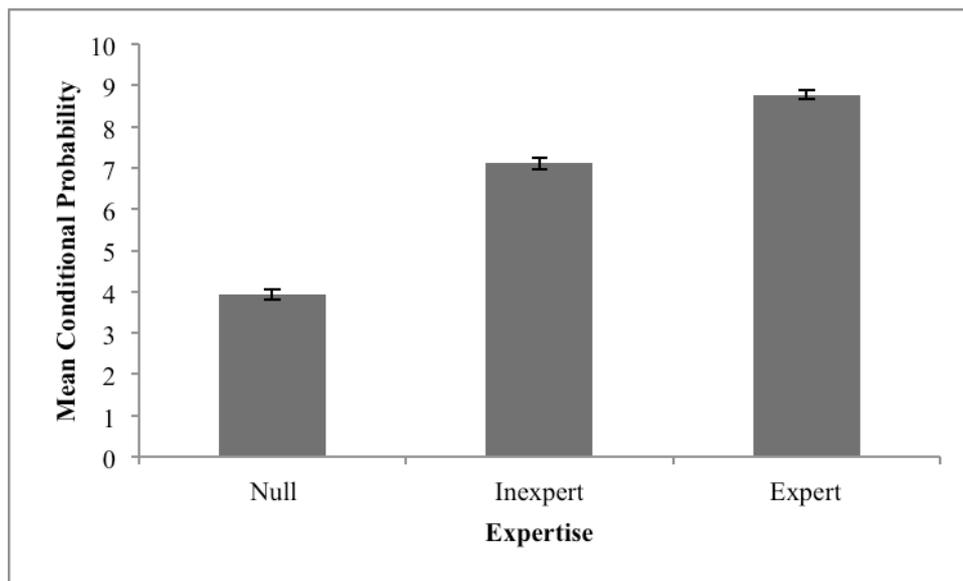


Figure 2.12. Mean Conditional Probability by condition; error bars are standard error. The data suggest a linear trend from the null condition ($M = 3.94$) to the inexpert condition ($M = 7.11$) and the expert condition ($M = 8.77$). Including Expertise significantly improved fit over the null model, $\chi^2(2) = 20.39$, $p < .001$. Table 2.14 reports the estimates of the fixed effects.

Table 2.14. Fixed effects of Expertise on Conditional Probability

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 3.94$	3.21, 4.67
Inexpert	$b = 3.17$	2.19, 4.15
Expert	$b = 4.83$	3.97, 5.69

Pairwise comparisons were performed on the least-squares means, using the Tukey correction for multiple comparisons. The increase from the null to inexpert

conditions was significant, $M = 3.17$, 95% CI [1.74, 4.60], $t(10.55) = 6.02$, $p < .001$.

The increase from the null to expert conditions was significant, $M = 4.83$, 95% CI [3.63, 6.04], $t(14.33) = 10.47$, $p < .001$. The increase from the inexpert to expert conditions was also significant, $M = 1.66$, 95% CI [.52, 2.81], $t(16.59) = 3.72$, $p = .005$.

The parameter estimates for the inexpert condition are consistent between the original and replication studies: for the original study, $M_{Inexpert} = 7.21$, 95% CI [6.58, 7.84]; and for the replication, $M_{Inexpert} = 7.11$, 95% CI [6.13, 8.09]. Once again, the estimates are tightly clustered, and the confidence interval for the replication includes that for the original study. The parameter estimates for the expert condition are also consistent between the original and replication studies: for the original study, $M_{Expert} = 8.59$, 95% CI [7.81, 9.36]; and for the replication study, $M_{Expert} = 8.77$, 95% CI [7.91, 9.63]. The estimates are close together, and there is considerable overlap in the confidence intervals (1.45).

The original and replication studies offer consistent data: the means are close together for all conditions and dependent variables, and the confidence intervals overlap substantially. The data suggest that Expertise has a focused effect on the conditional probability. There is no evidence of a reliable effect on the probability of the antecedent and the probability of the consequent.

2.6 General Discussion

The four experiments reported above offer insight into how our beliefs change when we learn a testimonial conditional. Experiments 2.1 and 2.2 showed that assertion increases the conditional probability, though with a diminishing return: ratings were not higher in the multiple-assertion condition than in the single-assertion condition. Assertion did not reliably influence the probability of the

antecedent or the probability of the consequent. Experiments 2.3 and 2.4 showed that source expertise likewise increases the conditional probability, but here there was no diminishing return: ratings were reliably higher for expert than inexperienced sources.

Again, though, the effect was limited to the conditional probability: expertise did not reliably influence the probability of the antecedent and the probability of the consequent. These data can be taken to generalize the findings of Stevenson and Over (2001) to testimonial contexts and to a different manipulation, the Assertion manipulation.

The data are also consistent with a large body of work in the psychology of reasoning which argues that the meaning of the conditional is closely associated with the conditional probability. As we saw in the Introduction, one way to interpret this association is the suppositional theory of meaning. On this account, when we hear a conditional *If A, B*, '[we] suppose (assume, hypothesize) that *A*, and make a hypothetical judgment about *B*, under the supposition that *A*, in light of your other beliefs' (Edgington, 2014). This judgment amounts to the conditional probability (for extended discussion, see, e.g., Evans & Over, 2004). Here is how a suppositional theorist could account for the present data. Participants supposed that the antecedent was true, and judged the conditional probability. In making their judgments, participants in the assertion task simply drew on the fact of assertion; participants in the source-expertise task drew also on the information about the source. But why should participants increase²⁹ the conditional probability? Here, the suppositional theorist can invoke the pragmatic assumption: that asserting a

²⁹ Or, more strictly, why should participants in the assertion conditions give higher ratings than the control condition? I am assuming, here, that the between-subjects design approximates learning a conditional.

conditional implies high conditional probability. In this experiment, then, participants moved from a non-committal probability to a higher one. Participants did not, however, reliably modify their belief in the antecedent. Coherently, then, they also did not reliably modify their belief in the consequent. This picture is also consistent with an account in linguistics which takes conditionals to be defined by remoteness: that is, in uttering an indicative conditional, a speaker is not committing to the truth (or falsity) of the antecedent (Elder & Jaszczolt, 2016).

The data sit less well with two other prominent accounts. The first takes the indicative conditional to be the material conditional. Although this account has had prominent supporters (Grice, 1975; Jackson, 1979), it is not widely supported in contemporary philosophy (Bennett, 2003; Hartmann & Rafiee Rad, 2017). The material conditional does, however, seem to underpin the approach to the conditional of Mental Models Theory³⁰ (Johnson-Laird & Byrne, 2002). The material conditional has limited appeal in part because its account of learning a conditional has counterintuitive consequences (for discussion, see Hartmann & Rafiee Rad, 2017; Popper & Miller, 1983). Assuming the material conditional, recall, means assuming that the probability of the conditional ‘If P, Q’ is equivalent to the disjunction $P(\sim P \vee Q)$. Accordingly, learning a conditional can be construed as conditioning on the disjunction. This conditioning is compatible with an increase in the conditional probability, $P(Q|P)$. To illustrate, take the simple case in which conditioning means assigning $P = 1$ to the disjunction. The disjunction $(\sim P \vee Q)$

³⁰The Mental Models account is not clearly expressed. Although Johnson-Laird and Byrne (2002) try to distinguish their account from the material conditional, referring to possibilities rather than truth, it still seems to reduce to the material conditional, plus highly flexible semantic and pragmatic modulation (for discussion, see Krzyżanowska, Collins, & Hahn, 2017).

corresponds to four conjunctions. Cases (1) to (3) make the disjunction true; case (4) makes it false:

- (1) $\sim P \ \& \ Q$
- (2) $\sim P \ \& \ \sim Q$
- (3) $P \ \& \ Q$
- (4) $P \ \& \ \sim Q$

Conditioning on $(\sim P \vee Q)$ means eliminating case (4): there are no P cases which are not also Q cases. The conditional probability $P(Q|P)$, therefore, increases to 1. But, less happily, as long as two constraints hold – $0 < P < 1$ and $0 < Q < 1$ – when we learn a conditional we should decrease the probability P and increase the probability of Q (for the proofs, see Popper & Miller, 1983). These latter predictions are not only counterintuitive; they are not supported by the data above.

The second prominent account takes learning a conditional to be governed by an attempt to minimize the difference between the prior and posterior distributions, defined formally as the Kullback-Leibler divergence (Hartmann & Rafiee Rad, 2017). This account has the same consequences for the probability of the antecedent as the material conditional as long as we assume a model for the conditional such as Figure 2.13, where ‘H’ represents the antecedent and ‘E’ the consequent:

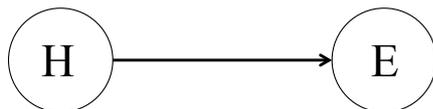


Figure 2.13. Simple Bayesian belief network for a conditional

Again, the data above do not support this downward revision in the probability of the antecedent. This problem can be avoided by building a more extensive causal model

RATIONALITY, PRAGMATICS, AND SOURCES

of the relevant situation. Doing so, as Hartmann and Rafiee Rad (2017) show, leads to other, more intuitively appealing revisions of belief. This strategy does not seem plausible for the present materials, however, as they are deliberately light on causal information. An alternative but related strategy, which models information about the sources, will be described in Chapter 4.

Although the present data cohere well with the dominant theory of the conditional within psychology and philosophy, there is reason to doubt how well the data will generalize. Consider the following intuitions.

Intuition 1: It is a sunny day, without a cloud in the sky, but you have not seen the weather forecast. Someone says to you, ‘If it rains this afternoon, they’ll have to postpone the tennis match’.

In this context, it seems likely that the hearer would increase their judgment of the probability of rain.

Intuition 2: You have been promised a job, and you believe that you are certain to get it. A trusted and knowledgeable colleague says to you, ‘If you get the job, we can collaborate more’.

In this context, it seems likely that the hearer would decrease their judgment of the probability of getting the job. This intuition may be clearer with the intonation ‘IF you get the job, we can collaborate more’. The relevant factor seems to be the hearer’s prior beliefs. Similar intuitions underlie the examples in Douven’s (2012) influential paper and the following passage from Evans and Over (2004, pp. 144-5):

“‘If p then q’ is not assertable – or at least has low relevance – in most contexts if $P(p)$ is too or if $P(q|p)$ is too low. As conditionals only apply to p-states, such states must normally be reasonably probable (at least in the near future for a conditional statement to have relevance.’

RATIONALITY, PRAGMATICS, AND SOURCES

Presumably, then, the hearer of a conditional will assume that the antecedent is reasonably probable. Chapter 4 will explore these intuitions further.

Finally, in this chapter, there is an important limitation in the present design: namely, that participants responded with point estimates. While such responses are simple and intuitive, they may conceal subtle belief change. For instance, participants' beliefs might be better captured by a distribution, the central tendency corresponding to the point estimate. The point estimates may not pick up changes to the underlying distributions. This possibility will be explored in the next chapter.

3 Conditionals, Testimony, and Interval Estimates

The previous chapter introduced novel data on how people's beliefs change when they learn a conditional. The data suggested that belief change focuses on the conditional probability, leaving the probability of the antecedent and of the consequent unaffected. The data sit well with the suppositional theory of the conditional, but they are intuitively surprising. Given their surprisingness and their relevance to theories of the conditional, it is worth probing deeper to ascertain whether belief change really is so focused.

This chapter explores the simplest adaptation to the experimental design: namely, to allow participants to respond by indicating a range of values. This adaptation allows a straightforward test of the possibility that there is belief change hiding behind the point estimates. Participants indicated a range by positioning two endpoints of a slider. This dependent measure should detect changes to an underlying probability distribution which would not have been detected by point estimates. The experiments below replicate those in the preceding chapter with this new dependent measure.

3.1 Experiment 3.1: Assertion and Interval Estimates

3.1.1 Method

Participants. 187 (92 female, 1 non-binary; average age 37.09 years) participants completed this experiment; 5 participants had already been excluded since they were non-native English speakers. As above, participants were recruited via the intermediary MTurk Data (www.mturkdata.com). High qualifications were set for the task to improve the quality of the data and maximize the number of native English speakers: participants had to be resident in the US, and have an overall

RATIONALITY, PRAGMATICS, AND SOURCES

approval rating of 99% for 1,000 previously completed tasks. Participants received a small fee, chosen to exceed the US minimum wage per minute.

Materials. These were the same as in Experiment 2.2.

Procedure. As in Experiments 2.1 and 2.2, the design manipulated Assertion (Null, Single, Multiple), and the dependent measures were the probability of the antecedent, the probability of the consequent, and the conditional probability. The design was between-subjects. Participants first gave informed consent, and were then assigned, in a round-robin fashion, to a condition. They were then shown the same general instructions as for Experiments 2.1 to 2.4. Since the dependent measure – ratings with a slider – was less transparent, participants were shown an additional set of instructions as follows:

We would like you to estimate the probability of various events. To do this, you will need to use a slider on the scale provided. Here is how you can use the slider. Imagine that you've been given the task of estimating the probability of heads when tossing a coin. You have tested the coin to your satisfaction, and you believe that the coin is fair. In this case, you might be pretty confident in estimating that the probability of heads is 50%. You can position the sliders in the following way for such an estimate.

Figure 3.1 shows the slider that participants saw at this point.

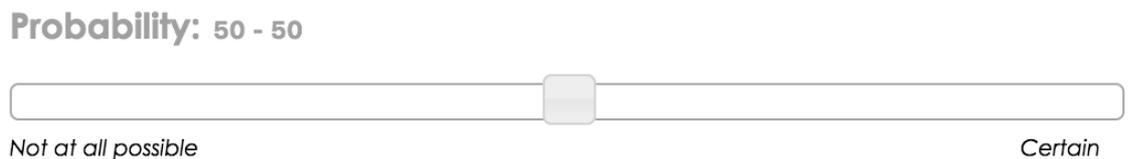


Figure 3.1. Illustrative slider for exact estimate of $P=.5$

The instructions then continued as follows:

RATIONALITY, PRAGMATICS, AND SOURCES

Now imagine that you have to estimate the probability of heads when tossing another coin. This time, the coin looks strange, and you believe it might be a coin from a magic shop. In this case, 50% might still be your best estimate.

But you might be less confident in your estimate. You can position the sliders in the following way for such an estimate.

Figure 3.2 shows the slider that participants saw at this point.



Figure 3.2. Illustrative slider for uncertain estimate

The instructions finished as follows:

In the following pages, you'll be asked to make a series of probability estimates. Your estimate can be anywhere on a scale from 0% to 100%. You may also want to choose an interval to show your confidence your estimates in the way described above.

Experiments 2.1 and 2.2 used different approaches for presenting the probability questions: Experiment 2.1 presented each probability question on a separate screen, so that participants saw each item three times but with different probability questions; Experiment 2.2 presented all the probability questions together. Other things being equal, the latter method is preferable. However, the latter method resulted in considerable visual complexity in the present experiment. Since the results did not differ reliably for Experiments 2.1. and 2.2, the present experiment presented the probability questions separately to minimize the visual complexity. After the probability question, participants were asked to 'Please use the sliders to give your personal estimate between 0% (not at all possible) and 100% (certain).'

The order of items was randomized. Finally, participants received debriefing information.

3.1.2 Results and Discussion

The analysis, here, takes the same general approach as with the point estimates: crossed random-effects models followed up by pairwise comparisons where appropriate. The analysis treated the dependent measure in two different ways. Firstly, the end points of the sliders were averaged to produce a point value (henceforth, ‘point values’) and entered into the models; this method allows replication of the data in Experiments 2.1 and 2.2. Secondly, the ranges of the sliders (henceforth, ‘slider ranges’) were entered into the models. Since these are two ways of looking at the same data, it is appropriate to correct for multiple analyses. Thus, the significance level is $p = .025$ below.

Probability of the Antecedent.

Point Values. Figure 3.3 shows the descriptive data for the effect of Assertion on the point values for the probability of the antecedent.

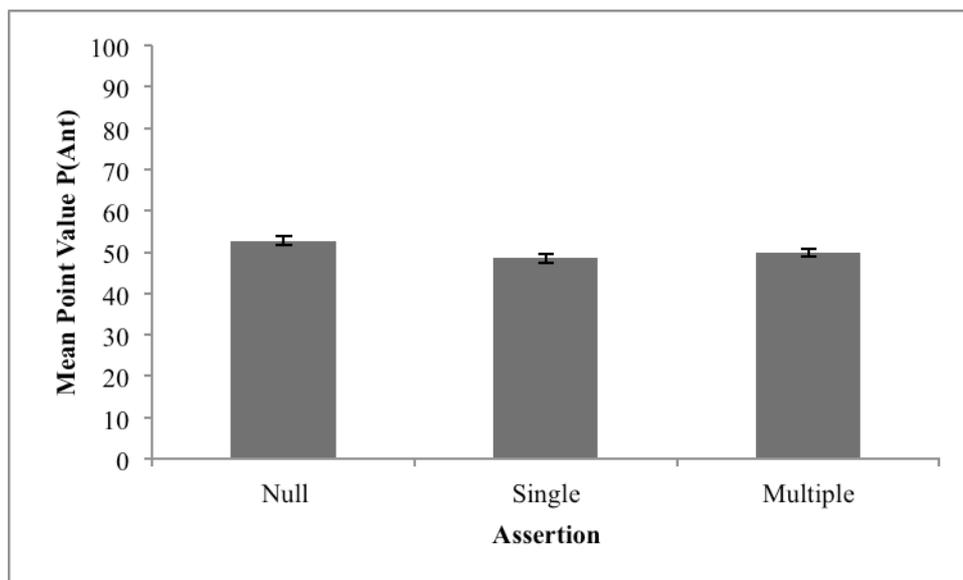


Figure 3.3. Mean point value of P(Antecedent) by condition; error bars are standard error

The descriptive data suggest small differences between conditions: there is a slight decrease from the null condition ($M = 52.61$) to the single condition ($M = 48.58$) and then a slight decrease to the multiple condition ($M = 49.94$). Including Assertion did not significantly improve fit over the null model, $\chi^2(2) = 1.49, p = .47$. Table 3.1 reports the estimates of the fixed effects.

Table 3.1. Fixed effects of Assertion on point values of P(Antecedent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 52.68$	45.19, 60.16
Single	$b = -4.10$	-10.56, 2.36
Multiple	$b = -2.74$	-9.12, 3.64

These data offer no clear evidence for an effect of Assertion, and are consistent with the data in the studies with point estimates.

Slider Ranges. Figure 3.4 shows the descriptive data for the effect of Assertion on the slider ranges.

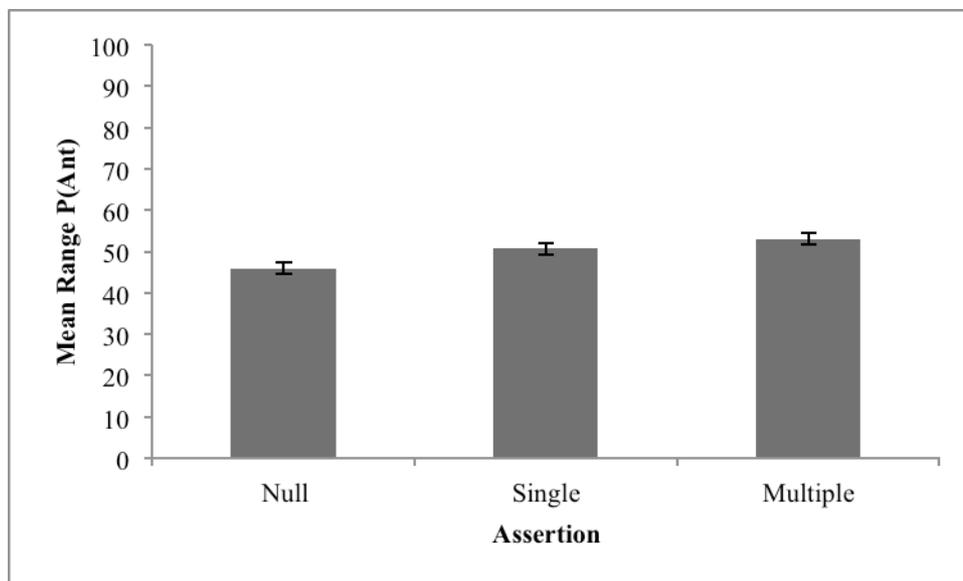


Figure 3.4. Mean slider range by condition; error bars are standard error

These data suggest that participants are uncertain about their probability estimates.

The data also slight trend for an increase in the slider ranges as the level of Assertion increases: from the null condition ($M = 45.95$) to the single condition ($M = 50.68$)

and the multiple condition ($M = 53.09$). However, including Assertion did not significantly improve fit over the null model, $\chi^2(2) = 3.08, p = .21$. Because of convergence problems, no estimate was included for the covariance of random intercepts and slopes. Table 3.2 reports the estimates for the fixed effects.

Table 3.2. Fixed effects of Assertion on slider ranges of P(Antecedent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 45.85$	40.06, 51.65
Single	$b = 4.83$	-3.37, 13.03
Multiple	$b = 7.24$	-1.02, 15.51

These data are consistent with those of Experiments 2.1 and 2.2 and with the analysis of the point values above.

Probability of the Consequent.

Point Values. Figure 3.5 shows the descriptive data for the effect of Expertise on the point values for the probability of the consequent.

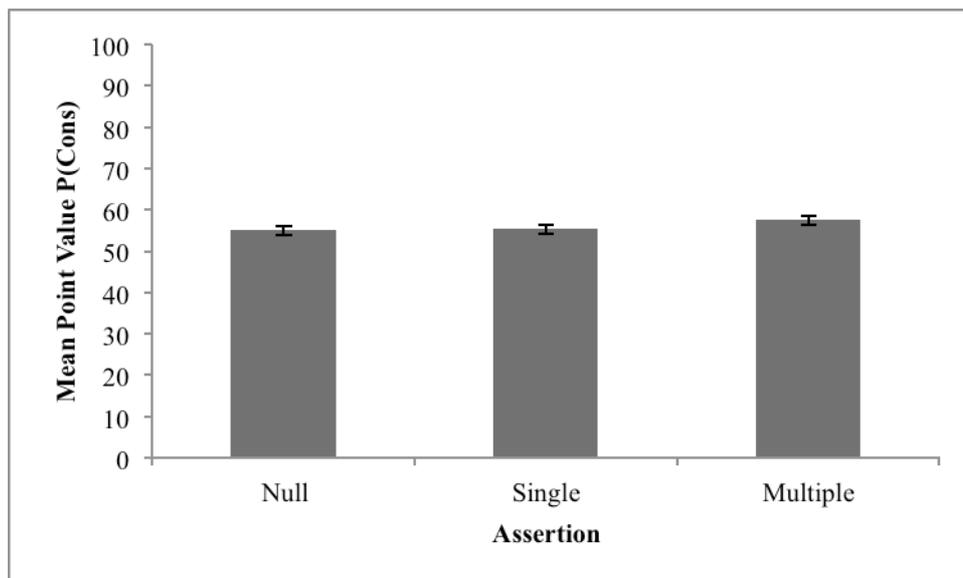


Figure 3.5. Mean point values of P(Consequent) by condition; error bars are standard error

RATIONALITY, PRAGMATICS, AND SOURCES

These data suggest, once again, that there is little evidence for an effect, as the means of the null condition ($M = 54.96$), the single condition ($M = 55.32$) and multiple condition ($M = 57.55$) are close together (though increasing). Including Assertion did not significantly improve fit over the null model, $\chi^2(2) = .10, p = .95$.

Table 3.3 report the fixed effects.

Table 3.3. Fixed effects of Assertion on point values of P(Consequent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 55.07$	47.80, 62.34
Single	$b = .25$	-10.56, 11.06
Multiple	$b = 1.05$	-10.57, 12.68

Slider Ranges. Figure 3.6 shows the descriptive data for the effect of Assertion on the slider ranges for the probability of the consequent.

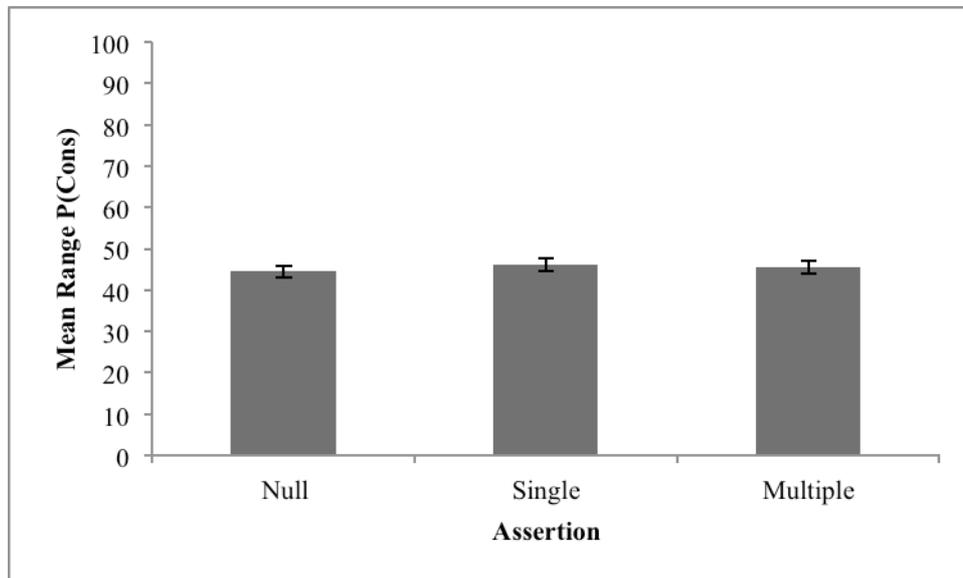


Figure 3.6. Mean range of P(Consequent) by condition; error bars are standard error. The data suggest little evidence for an effect: ranges increase slightly from the null condition ($M = 44.56$) to the single condition ($M = 46.20$) and fall again to the multiple condition ($M = 45.48$). Including Assertion did not significantly improve fit

over the null model, $\chi^2(2) = .45, p = .80$. Table 3.4 reports the estimates of the fixed effects.

Table 3.4. Fixed effects of Assertion on slider ranges for P(Consequent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 44.42$	38.51, 50.32
Single	$b = 1.78$	-7.43, 10.99
Multiple	$b = 3.43$	-6.55, 13.41

Taken together the data for P(Consequent) are consistent with the data for Experiments 2.1 and 2.2: there is no reliable change in the probability of the consequent, whether it is measured as a point estimate or as an interval.

Conditional Probability.

Point Values. Figure 3.7 shows the descriptive data for the effect of Assertion on the point values for the conditional probability.

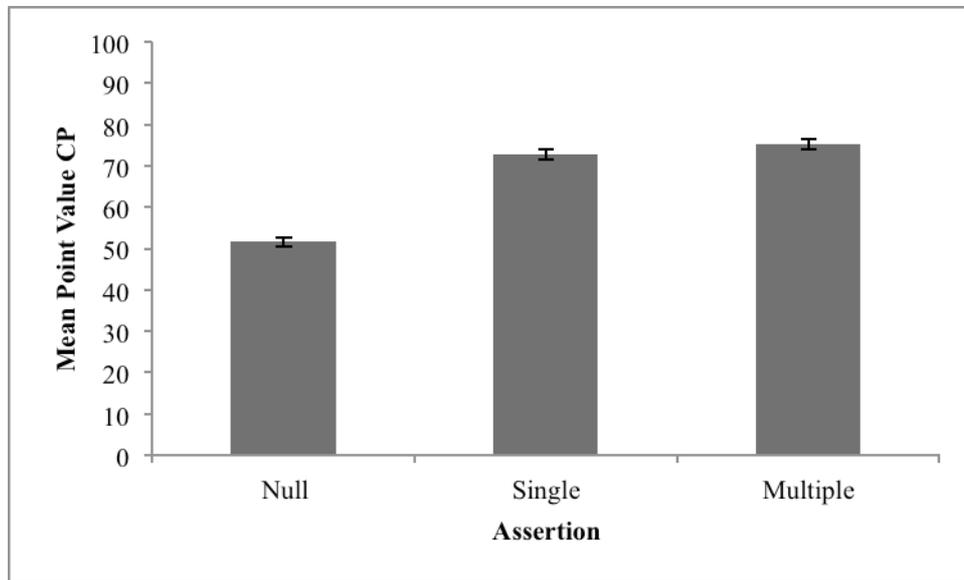


Figure 3.7. Mean point values of Conditional Probability by condition; error bars are standard error

The data suggest a linear trend for Assertion to increase the slider averages for the conditional probability. There is an increase from the null condition ($M = 51.62$) to the single condition ($M = 72.62$) and a smaller increase to the multiple condition (M

RATIONALITY, PRAGMATICS, AND SOURCES

= 75.19). Indeed, including Assertion significantly improved fit over the null model, $\chi^2(2) = 17.76, p < .001$. Table 3.5 reports the estimates of the fixed effects.

Table 3.5. Fixed effects of Assertion on Conditional Probability

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 51.34$	45.09, 57.59
Single	$b = 21.28$	13.53, 29.04
Multiple	$b = 23.74$	15.21, 32.27

Pairwise comparisons were performed on the least-squares means using the Tukey correction for multiple comparisons. The increase from null to single assertion was significant, $M = 21.28$, 95% CI [11.67, 30.90], $t(72.70) = 5.30, p < .001$. So too was the increase from null to multiple assertion, $M = 23.74$, 95% CI [12.91, 34.57], $t(37.93) = 5.35, p < .001$. The increase from single to multiple assertion, in contrast, was not significant, $M = 2.45$, 95% CI [-6.32, 11.22], $t(148.91) = .66, p = .79$. These data are consistent with the data from the point-estimate studies. There is, again, evidence for a plateauing in the effect of Assertion.

Slider Ranges. Figure 3.8 shows the descriptive data for the effect of Assertion on the slider ranges for the conditional probability.

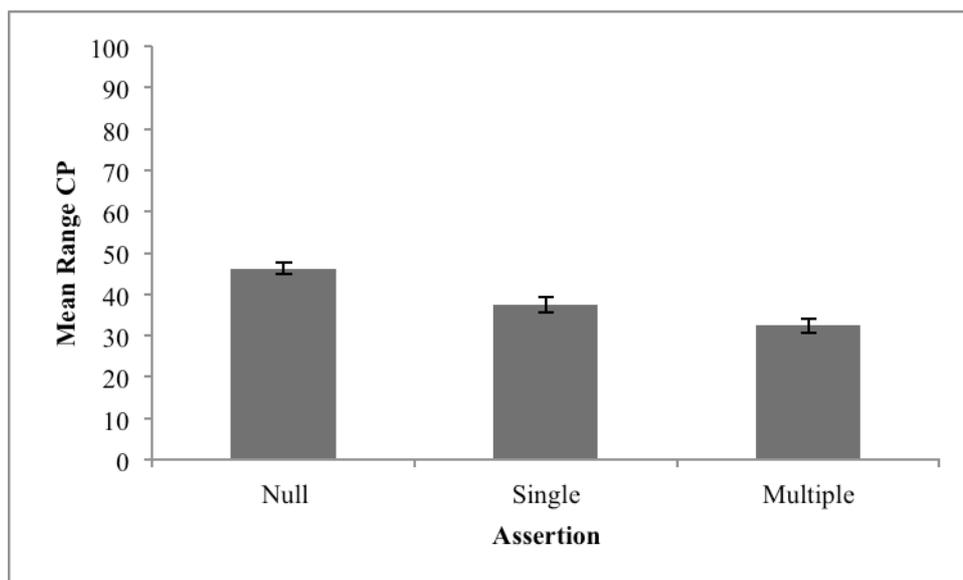


Figure 3.8. Mean range of Conditional Probability by condition; error bars are standard error

RATIONALITY, PRAGMATICS, AND SOURCES

There is evidence for a linear trend of Assertion: ranges decrease from the null condition ($M = 46.18$) to the single condition ($M = 37.60$) and the multiple condition ($M = 32.41$). Because of convergence problems, the mixed-effects, here, excluded the estimates of the covariance for random slopes and intercepts. Including Assertion improved fit over the null model. We have adopted the more stringent significance level of $p = .025$. The model improved fit, $\chi^2(2) = 7.26, p = .027$. Given this closeness, further analyses are justified.

Table 3.6. Fixed effects of Assertion on ranges for Conditional Probability

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 46.42$	39.44, 53.41
Single	$b = -8.83$	-18.95, 1.30
Multiple	$b = -13.85$	-24.02, -3.68

Pairwise comparisons were performed on the least-squares means, using the Tukey correction for multiple comparisons. The decrease from the null to single conditions was not significant, $M = -8.83$, 95% CI [-21.13, 3.48], $t(170.09) = -1.70, p = .21$. The decrease from the null to multiple conditions was significant, $M = -13.85$, 95% CI [-26.21, -1.49], $t(171.43) = -2.65, p = .02$. The decrease from the single to multiple conditions was not significant, $M = -5.02$, 95% CI [-5.02, 5.38], $t(166.71) = -.93, p = .62$. It is well to note that, since participants made their response on a bounded scale, there will inevitably be some shortening of the ranges at the ends of the scale. These data should, therefore, be interpreted cautiously. However, these data can be taken to suggest that Assertion decreases the slider range: that people's beliefs become more precise as the level of Assertion increases.

3.2 Experiment 3.2: Expertise and Interval Estimates

3.2.1 Method

Participants. 173 (52 female, 3 non-binary; average age 38.43 years) participants completed the task; 5 participants had already been excluded since they were non-native speakers of English. As above, participants were recruited via the intermediary MTurk Data (www.mturkdata.com). High qualifications were set for the task to improve the quality of the data and maximize the number of native English speakers: participants had to be resident in the US, and have an overall approval rating of 99% for 1,000 previously completed tasks. Participants received a small fee, chosen to exceed the US minimum wage per minute.

Materials. These were the same as for Experiment 2.4.

Procedure. This experiment manipulated Expertise (Null, Inexpert, Expert). The procedure was identical to that of Experiment 3.1.

3.2.2 Results & Discussion

As with the Assertion data, this section adopts a more stringent significance level of $p = .025$ to account for multiple analyses of the same data.

Probability of the Antecedent.

Point Values. Figure 3.9 shows the descriptive data for the effect of Expertise on the slider averages for the probability of the antecedent.

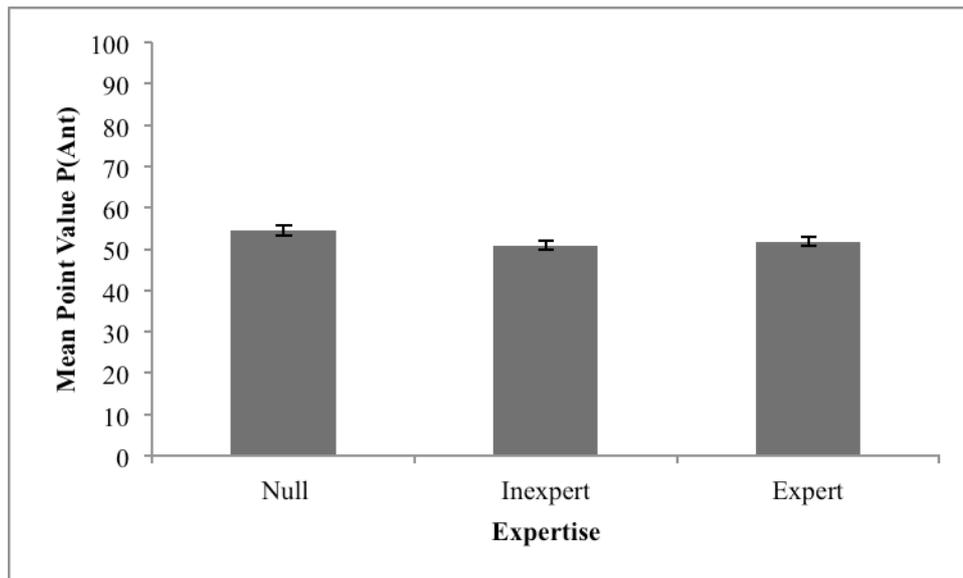


Figure 3.9. Mean point values of P(Antecedent) by condition; error bars are standard error. The data do not suggest a clear trend: there is a small decrease from the null condition ($M = 54.58$) to the inexpert condition ($M = 50.86$) and then an increase to the expert condition ($M = 51.71$). Including Expertise did not significantly improve fit over the null model, $\chi^2(2) = 1.88, p = .39$. Table 3.7 reports the estimates of the fixed effects.

Table 3.7. Fixed effects of Expertise on point values of P(Antecedent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 54.63$	48.86, 60.40
Inexpert	$b = -3.84$	-9.57, 1.89
Expert	$b = -2.93$	-8.57, 2.70

Slider Ranges. Figure 3.10 shows the descriptive data for the effect of Expertise on the slider ranges for the probability of the antecedent.

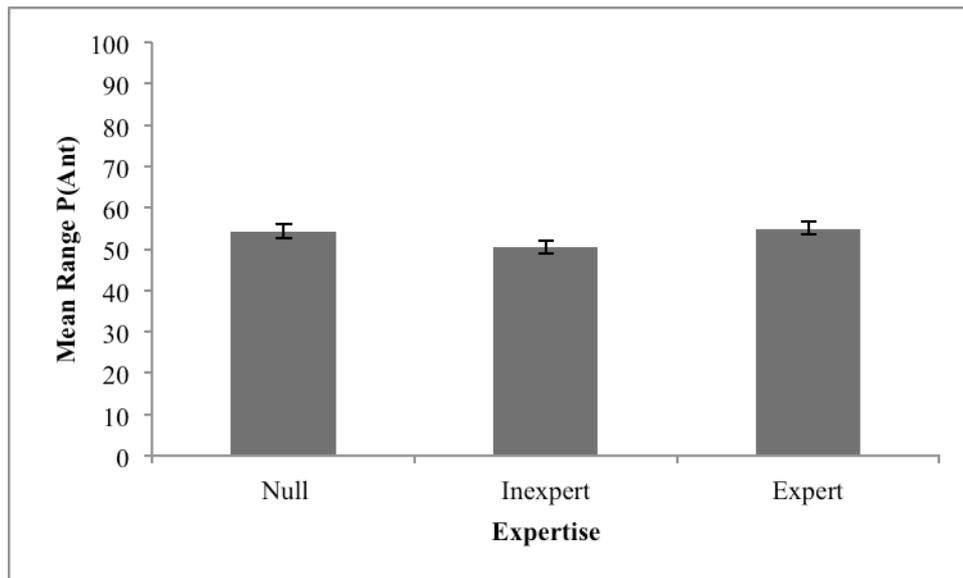


Figure 3.10. Mean slider ranges for P(Antecedent) by condition; error bars are standard error. As above, there is little evidence for a trend: the means for the null condition ($M = 54.27$), inexpert condition ($M = 50.47$) and expert condition ($M = 54.94$) are clustered, with the inexpert condition slightly lower. Because of convergence problems, the models, here, exclude the random intercepts (though not the random slopes) for topic. Including Expertise did not significantly improve fit over the null model, $\chi^2(2) = 1.25, p = .53$.

Table 3.8. Fixed effects of Expertise on slider ranges for P(Antecedent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 54.23$	46.93, 61.54
Inexpert	$b = -3.86$	-12.45, 4.73
Expert	$b = .74$	-8.27, 9.76

Probability of the Consequent.

Point values. Figure 3.11 shows the descriptive data for the effect of Expertise on the point values for the probability of the antecedent.

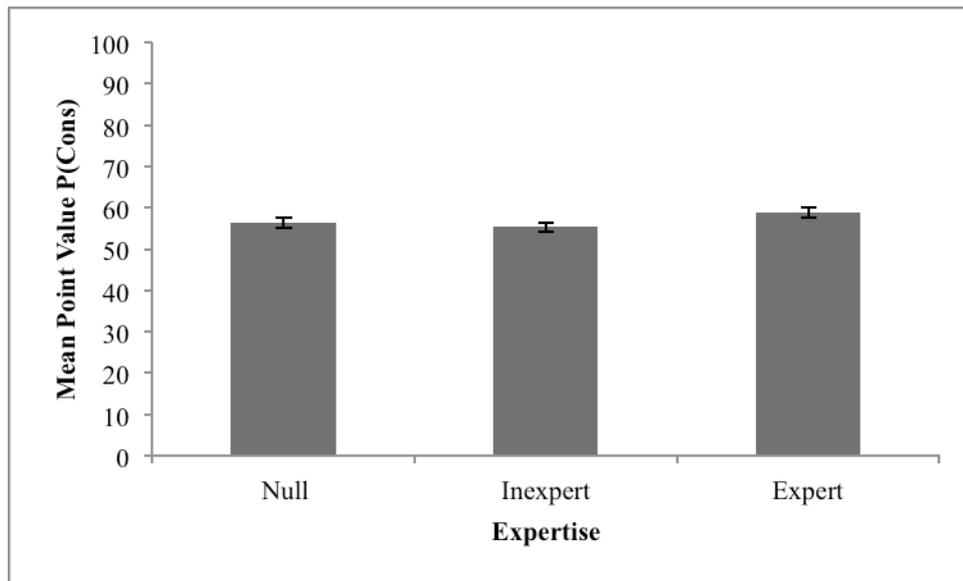


Figure 3.11. Mean point values of P(Consequent) by condition; error bars are standard error. The means are, again, grouped close together for the null condition ($M = 56.33$), inexpert condition ($M = 55.31$) and expert condition ($M = 58.91$). Including Expertise did not significantly improve fit over the null model, $\chi^2(2) = 1.27, p = .53$.

Table 3.9. Fixed effects of Expertise on point values for P(Consequent)

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 56.38$	51.03. 61.72
Inexpert	$b = -1.11$	-8.38. 6.15
Expert	$b = .253$	-8.84, 13.90

Slider Ranges. Figure 3.12 shows the descriptive data for the effect of Expertise on the probability of the consequent.

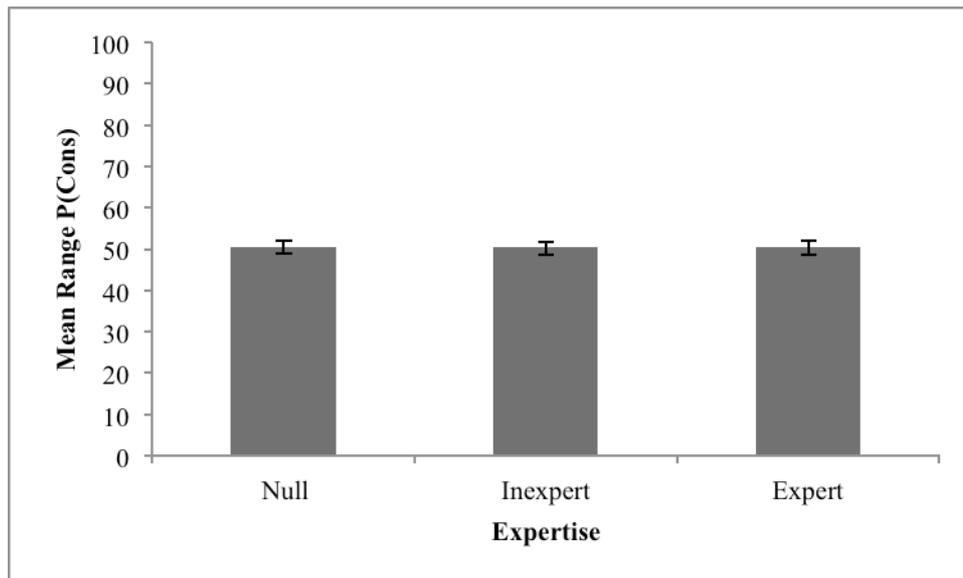


Figure 3.12. Mean slider ranges for P(Consequent) by condition; error bars are standard error

Once again, the means are tightly clustered for the null condition ($M = 50.47$), the inexpert condition ($M = 50.32$) and the expert condition ($M = 50.34$). The models exclude the random intercept of items because of convergence problems. Including Expertise did not significantly improve fit over the null model, $\chi^2(2) = .0008, p = 1$.

Table 3.10. Fixed effects of Expertise on slider ranges for P(Consequent); error bars are standard error

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 50.40$	46.93, 61.54
Inexpert	$b = -.12$	-12.45, 4.73
Expert	$b = -.04$	-8.27, 9.76

Conditional Probability.

Point Values. Figure 3.13 shows the descriptive data for the effect of Expertise on the point values for the Conditional Probability.

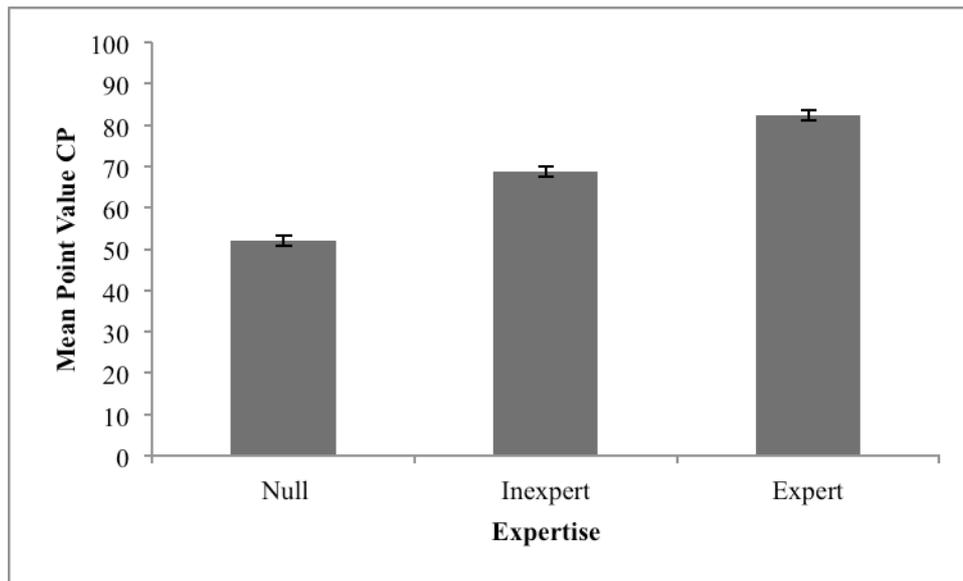


Figure 3.13. Mean point values for Conditional Probability by condition; error bars are standard error

The data suggest a linear trend for the slider averages to increase with Expertise from the null condition ($M = 52.02$) to the inexpert condition ($M = 68.68$) and the expert condition ($M = 82.26$). Including Expertise significantly improved fit over the null model, $\chi^2(2) = 23.84, p < .001$. Table 3.11 reports the estimates of the fixed effects.

Table 3.11. Fixed effects of Expertise on point values for Conditional Probability

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 52.05$	47.18, 56.91
Inexpert	$b = 16.61$	9.87, 23.35
Expert	$b = 30.22$	23.41, 37.04

Pairwise comparisons were performed on the least-square means, using the Tukey correction for multiple comparisons. The increase from null to inexpert conditions was significant, $M = 16.61$, 95% CI [8.29, 24.94], $t(75.78) = 4.77, p < .001$. So too was the increase from null to expert conditions, $M = 30.22$, 95% CI [21.82, 38.62], $t(81.95) = 8.59, p < .001$. The increase from inexpert to expert conditions was also significant, $M = 13.61$, 95% CI [5.00, 22.21], $t(67.75) = 3.79, p < .001$. These data replicate the findings of the point-estimate studies.

Slider Ranges. Figure 3.14 shows the descriptive data for the effect of Expertise on the slider ranges for the conditional probability.

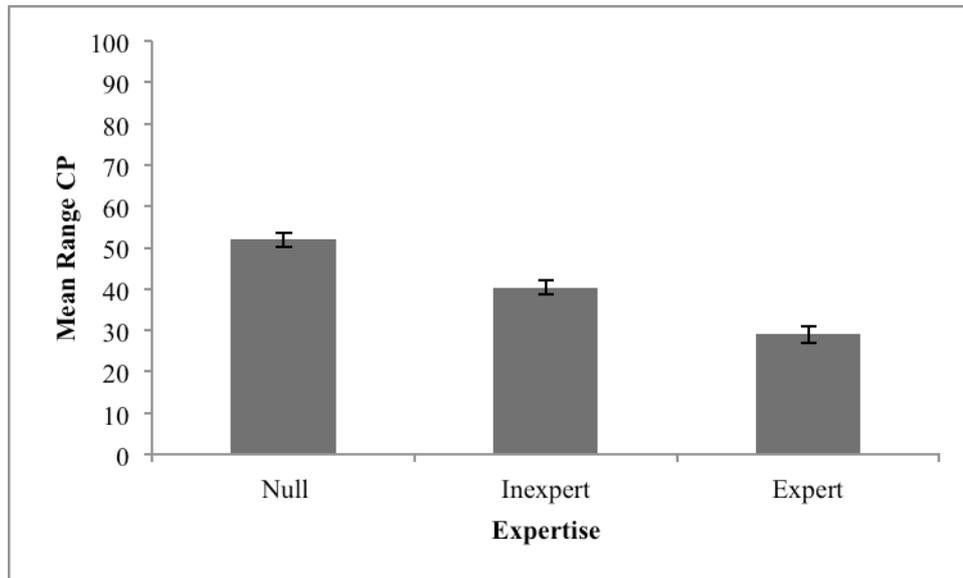


Figure 3.14. Mean slider ranges for Conditional Probability by condition; error bars are standard error

The data suggest a linear trend for the slider ranges to decrease with Expertise, from the null condition ($M = 51.93$) to the inexpert condition ($M = 40.38$) and the expert condition ($M = 28.97$). The models, here, excluded the estimate of the covariance of random slopes and intercepts. Including Expertise significantly improved fit over the null condition, $\chi^2(2) = 16.84, p < .001$. Table 3.12 reports the estimates of the fixed effects.

Table 3.12. Fixed effects of Expertise on slider ranges for Conditional Probability

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 51.93$	44.31, 59.55
Inexpert	$b = -11.55$	-22.47, -.63
Expert	$b = -22.98$	-33.66, -12.30

Pairwise comparisons were performed on the least-squares means, using the Tukey correction for multiple comparisons. The decrease from null to inexpert conditions was non-significant, $M = -11.55$, 95% CI [-24.91, 1.80], $t(117.56) = -2.05, p = .10$.

The decrease from null to expert conditions was significant, $M = -22.98$, 95% CI [-35.99, -9.97], $t(152.25) = -4.18$, $p < .001$. The decrease from inexpert to expert conditions was non-significant, $M = -11.42$, 95% CI [-11.42, 5.63], $t(122.69) = -2.03$, $p = .11$. As with the Assertion data, there is a trend for participants' beliefs to become more precise; but, again as with the Assertion, this increasing precision may partly be due to inevitable narrowing of the range towards the end of the bounded scale.

3.3 General Discussion

This chapter has investigated belief change from learning a conditional as measured by interval estimates. The data support the findings of Chapter 2. On learning a conditional, participants did not reliably adjust their judgments of the probability of the antecedent or the probability of the consequent, whether the estimates were analysed as point estimates or intervals. For these probabilities, their estimates were wide, indicating high uncertainty. However, participants did, again, reliably increase their judgments of the conditional probability; they also appeared to make their judgments more precise.

These data arose from considering the conditional in simple assertions, that is, in simple pragmatic contexts. The data offer suggestive evidence about how people change their beliefs when they learn a conditional from a speaker; hence, about the information content – the meaning (pragmatics included) – of the conditional. All the experiments, thus far, have derived predictions from rational models in the philosophy of testimony, illustrating the value of considering pragmatics, sources, and rationality together.

Experiments 3.1 and 3.2 offer further evidence against the material conditional and its psychological variants. There was, again, no good evidence that people decreased their belief in the probability of the antecedent or that they increased their belief in the probability of the consequent. The data also count as further evidence against the simple Bayesian belief network with minimization of the Kullback-Leibler Divergence. This account also predicts that people revise their belief in the antecedent downwards. Experiments 3.1 and 3.2 suggest, then, that the data in Chapter 2 were not just an artefact of the response scale.

In contrast, these data are consistent with two competitors of the material conditional. As discussed in Chapter 2, the first of these theories, the suppositional theory, sits well with focal change to the conditional probability. Recall that, on this account, people suppose the antecedent to be true, and then judge the conditional probability of the consequent given the antecedent. The theory does not imply change to the probability of the antecedent or the probability of the consequent. Although the theory does not seem to imply that interval estimates should decrease, it is not inconsistent with this finding. Similarly, the data are consistent with the idea of remoteness: that is, the idea that the speaker of a conditional makes no commitment to the truth or falsity of the antecedent of an indicative conditional.

In sum, the data so far support the leading theory of the conditional in psychology, namely the suppositional theory, and they extend the data of Stevenson and Over (2001). But there remain issues to be resolved. Firstly, each Assertion experiment has suggested a diminishing return, or plateauing: there is a reliable increase from the null to single-assertion conditions, but no reliable difference between single and multiple-assertion conditions. This finding is intuitively surprising and, potentially, problematic from the point of view of rationality. In

RATIONALITY, PRAGMATICS, AND SOURCES

many (though not all) circumstances, a good case can be made that a hearer should put more weight in testimony from multiple sources than from a single source (Bovens & Hartmann, 2003). Secondly, the tension remains between the finding of no reliable change to the probability of the antecedent or the probability of the consequent and the intuition that the probability of the antecedent, at least, will shift in some contexts, as discussed in Chapter 2.

The next chapter will explore the issues raised in the preceding paragraph. In so doing, it will reveal findings which illustrate the advantages of considering sources, rationality, and pragmatics together. The chapter will also discuss a model for capturing the effects revealed by these three chapters on testimony and the conditional.

4 Independent Testimony, Priors, and a Model of Testimonial Conditionals

This chapter rounds off the section of the thesis on conditionals and testimony. It will address two outstanding issues from previous chapters - the role of multiple sources, and the role of the prior – and report experiments on them. The chapter will close by drawing together the data from these three chapters in a general discussion. Taken as a whole, the data demonstrate the limits of current modelling techniques. This point will be illustrated by discussing attempts to model the data with Bayesian belief networks (Collins, Krzyżanowska, Wheeler, Hartmann, & Hahn, 2017).

4.1.1 Multiple Testimony

Chapters 2 and 3 reported a consistent difference between the assertion and expertise experiments, which this chapter will explore. In each assertion experiment, the Assertion variable increased ratings of the conditional probability, but this effect plateaued. In no assertion experiment was there a reliable difference between the single and multiple conditions. In each expertise experiment, however, the Expertise variable increased ratings of the conditional probability at each level: inexpert sources reliably increased the conditional probability over the null condition; expert sources reliably increased the conditional probability over inexpert sources. In only one analysis was there no reliable difference between the inexpert and expert conditions: namely, when the sliders data, from Chapter 3, were treated as ranges (though when point values were calculated, there was a reliable difference).

This difference requires explanation. Intuitively, it is puzzling that participants were prepared to take one source at their word but that several sources made no reliable difference. Again intuitively, corroboration is desirable: think of

RATIONALITY, PRAGMATICS, AND SOURCES

multiple witnesses to a crime, or multiple sources for a news story (Sullivan, 2016). There are also formal results suggesting that multiplicity of sources can, in the right circumstances, be decisive (Bovens & Hartmann, 2003). Consequently, it is important to test whether the ineffectiveness of multiple testimonies is a robust finding or merely an artefact of the experimental materials.

It is worth examining these intuitions and formal results more closely. Take, first, the intuition that having more sources is better. This intuition will not always hold. In a court case, for instance, multiple witnesses are desirable, but the multiplicity is only probative if the witnesses have not conferred and agreed on story (Kadane & Schum, 1996). In other words, the witnesses need to be independent to some degree. The formal results, too, are subtle. It may sometimes be wise to place more trust in groups than individuals, since group accuracy often surpasses individual accuracy (Surowiecki, 2005). Indeed, as long as individual group members have an accuracy above $P = .5$ and make independent judgments, group accuracy converges on the truth (Condorcet, 1785; for discussion, see, e.g., Hahn, Harris, & Corner, 2016). Multiplicity can also be an asset in testimony. Here, at least three factors are relevant: the coherence of the claims; the number, or proportion, of sources making a claim; and the reliability of the sources (Bovens & Hartmann, 2003; for empirical evidence on coherence, see Harris & Hahn, 2009). Coherent evidence is often, though not always, more likely to be true (Bovens & Hartmann, 2003; Glass, 2007; Olsson, 2005), especially when the sources are independent (Olsson & Schubert, 2007; Schubert, 2012). The absolute number of sources is important when the sources are independent (List, 2004); the proportion of sources is important when the sources are dependent (Bovens & Hartmann, 2003). Lastly, reliability and perceived bias towards a positive report influence whether it is better

for testimony to be dependent or independent (Bovens & Hartmann, 2003; for discussion, see Hahn, Harris, & Corner, 2016).

How do these normative issues bear on the Assertion variable in the present studies? Take, for example, the following item from the experimental materials:

Adam, Barbara, Nick and Sue are at a large car dealership. They say, ‘If a car on this lot is a Mercedes, then it’s black.’

Here, the sources are clearly making a maximally coherent claim, but the relationship between the sources is ambiguous. The sources could be fully dependent, fully independent, or somewhere in between. If the sources are fully dependent, then there is little, if anything, to be gained in revising one’s belief higher than for a single source. Compare the existing item with the following novel item:

Adam, Barbara, Nick and Sue are each at a large car dealership. They don’t know each other, and haven’t spoken to each other. They each tell you independently, ‘If a car on this lot is a Mercedes, then it’s black.’

This example can only reasonably be read as independent testimony. In these simple cases, in which the only information is the claim and the number of people making the claim, it seems both likely that participants will, and desirable that they should, place more weight in multiple independent testimony than in single testimony (Hahn et al., 2016).

Experiment 4.1, below, uses such items to test the robustness of the plateauing effect: the ineffectiveness of the multiple assertion condition.

4.1.2 Priors

Chapters 2 and 3 have found that assertion and expertise predicted ratings of the conditional probability, but found no such evidence for an effect on the

probability of the antecedent or the probability of the consequent. These data seem intuitively surprising. Chapter 2 discussed the following two intuitions:

Intuition 1: It is a sunny day, without a cloud in the sky, but you have not see the weather forecast. Someone says to you, ‘If it rains this afternoon, they’ll have to postpone the tennis match’.

The hearer increases his/her judgment of the probability of rain.

Intuition 2: You have been promised a job, and you believe that you are certain to get it. A trusted and knowledgeable colleague says to you, ‘If you get the job, we can collaborate more’.

The hearer decreases his/her judgment of the probability of getting the job.

These intuitions can be buttressed by natural-language pragmatics. Take, first, Intuition 1. The conditional, here, would be an odd thing to say unless there were a considerable chance that the antecedent obtains: that there will be rain. This is a stronger phenomenon than the standard presupposition of the indicative conditional, that the antecedent is possible (see, e.g., Leahy, 2011). At work is an assumption that speakers should be relevant (e.g. Clark, 2013; Grice, 1975; Sperber & Wilson, 1995). Take, now, Intuition 2. Imagine a speaker who has to choose between two statements: a stronger statement, and a weaker statement. If the speaker is similarly confident in the stronger and weaker statements, then the speaker should choose the stronger statement (e.g. Grice, 1975; Jackson, 1979). A speaker who makes a weaker statement, but otherwise appears cooperative, may be taken to mean that the stronger statement does not hold (Grice, 1975; Horn, 1989). In this case, the speaker could have said ‘Since you will get the job, we can collaborate more’. The speaker’s choice of ‘if’ might, in the right context, be taken to mean that the speaker does not believe the hearer is certain to get the job.

Experiments 4.2 to 4.5 adapt the existing materials to explore the effects of prior probability on the probability of the antecedent and the probability of the consequent. Although the intuitions have only treated the probability of the antecedent, the probability of the consequent is included for completeness.

4.2 Experiment 4.1: Independent testimony

4.2.1 Method.

Participants. 240 participants (115 female, 1 other; average age 33.91 years) completed the task; 2 participants had previously been excluded since their first language was not English. Participants were recruited via the intermediary MTurk Data (www.mturkdata.com). High qualifications were set for the task to improve the quality of the data and maximize the number of native English speakers: participants had to be resident in the US, and have an overall approval rating of 99% for 1,000 previously completed tasks. Participants received a small fee, chosen to exceed the US minimum wage per minute.

Materials. The experimental items were the same as for Experiments 2.2 and 3.1. To realize the independent testimony condition, the existing multiple-assertion condition was adapted in the following way (changes in italics).

Adam, Barbara, Nick and Sue are *each* at a large car dealership. *They don't know each other, and haven't spoken to each other.* They *each* tell you *independently*, 'If a car on this lot is a Mercedes, then it's black.'

Since the additional text was exactly the same for each item, and the items were otherwise identical to those in Experiments 2.2 and 3.1, I will not report the full set of items here.

Procedure. The design manipulated Assertion (Null, Single, Multiple, Multiple Independent). Note that ‘Multiple’, here, refers to the original multiple condition, which is ambiguous. The task focused on the conditional probability $P(\text{Consequent}|\text{Antecedent})$, measured on a Likert-style scale from 0 (not at all possible) to 10 (certain). The task was between-subjects. After giving informed consent, participants were assigned, in a round-robin fashion, to a condition. Participants first read instructions, which were the same as the instructions for Experiment 2.2. Participants then read and rated 7 items, the order of presentation being randomized. The dependent measure was worded as in Experiment 2.2.

4.2.2 Results

Figure 4.1 shows the descriptive statistics for the effect of Assertion, including the new multiple independent assertion condition, on the conditional probability.

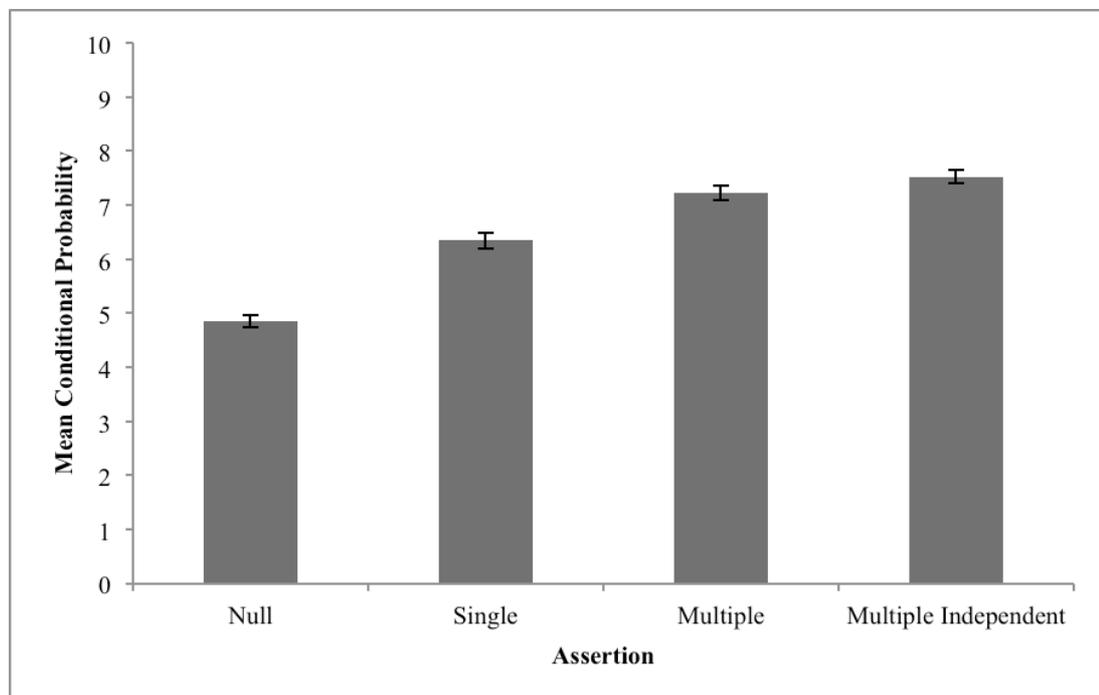


Figure 4.1. Mean Conditional Probability by condition; error bars are standard error

The data suggest a linear trend of Assertion, from the null condition ($M = 4.85$) to the single condition ($M = 6.35$), the multiple condition ($M = 7.22$), and the multiple

independent condition ($M = 7.52$). A mixed-effects model (with full random-effects structure) supported this picture. Including Assertion significantly improved fit over the null model, $\chi^2(3) = 14.55, p = .002$. Table 4.1 reports the fixed effects.

Table 4.1. Fixed effects of Assertion on Conditional Probability

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 4.85$	3.92, 5.79
Single	$b = 1.50$.43, 2.56
Multiple	$b = 2.37$	1.32, 3.42
Multiple Independent	$b = 2.67$	1.53, 3.82

Pairwise comparisons were performed using the Tukey correction for multiple comparisons. The increase from null to single conditions was not significant, though it approached significance, $M = 1.50$, 95% CI [-.09, 3.08], $t(18.95) = 2.66, p = .07$. The increase from null to multiple conditions was significant, $M = 2.37$, 95% CI [.80, 3.94], $t(18.16) = 18.16, p = .002$. So was the increase from null to multiple independent conditions, $M = 2.67$, 95% CI [.93, 4.42], $t(4.39) = .002$. The increase from single to multiple conditions was not significant, though it approached significance, $M = .87$, 95% CI [-.06, 1.81], $t(118.64) = 2.43, p = .08$. The increase from single to multiple independent conditions was significant, $M = 1.17$, 95% CI [.10, 2.24], $t(80.38) = 2.88, p = .03$. Lastly, the increase from multiple to multiple independent conditions was not significant, $M = .30$, 95% CI [-.75, 1.35], $t(83.02) = .75, p = .88$.

4.2.3 Discussion

Experiment 4.1 offers new data on the effect of multiple sources. For the first time, a condition with multiple sources is rated reliably higher than one with a single source. The new condition combines multiplicity with independence; this combination seems enough to tip the result over into significance. These data seem

to remove, therefore, a surprising finding: the plateauing effect in the Assertion variable. Moreover, participants seem to be sensitive to a normative factor, independence, in making their estimates of the conditional probability.

The above interpretation requires some caveats. Firstly, there is a difference between the present data and that for the previous assertion experiments: participants, here, rated the single-assertion condition somewhat lower than in previous experiments. Table 4.2 reports the fixed effects of single assertion for each of the three point-estimate studies.

Table 4.2. Fixed effects and confidence intervals of single assertion for the point-estimate studies

Study	Fixed Effect of Single Assertion	95% CI
Experiment 2.1	$b = 2.69$	1.32, 4.06
Experiment 2.2	$b = 2.54$	1.75, 3.34
Experiment 4.1	$b = 1.5$.43, 2.56

If we take Cumming's (2012) rule of thumb for the consistency of confidence intervals, then we calculate the average margin of error, here 1.07, and the overlap between confidence intervals, and judge whether the overlap is greater than half the length of the margin of error. Here, the smallest overlap is .81, which suggests, then, that the confidence intervals are consistent. However, the estimate for the multiple independent assertion condition also falls within, or overlaps considerably with, the confidence intervals for the fixed effect of single assertion for both Experiments 2.1 and 2.2. These considerations suggest an alternative explanation: the single and multiple conditions are marginally significantly different, and the single and multiple-independent conditions are significantly different, because the single condition is rated rather lower than in previous studies. These low ratings lack an

obvious explanation, since the materials (apart from the extra condition) and procedure were identical to Experiment 2.2.

A second caveat is this. As we saw earlier in this chapter, multiplicity is a separate factor from (in)dependence. The present design does not include an unambiguous multiple dependent condition. It cannot, therefore, directly address the question of whether participants are sensitive to the relationship between the sources. Moreover, the difference between single and multiple independent conditions, though significant, is not large. There is little space, then, for a significant difference from a hypothetical multiple dependent condition, unless that condition were treated as equivalent to single testimony.

We must draw cautious conclusions here. These data suggest that the Assertion variable does not necessarily lead to a plateauing effect: that this effect depends on how participants interpret the multiple-assertion condition. But more data should be collected, both to extend the present data and to explore, further, whether participants are sensitive to the crucial factors of multiplicity and (in)dependence of sources. Future studies could combine modelling of testimonial conditionals (on which, more later) and empirical work with additional variables, to ask whether models predict that multiplicity and independence interact in the complex ways suggested by Bayesian modelling of testimony (Bovens & Hartmann, 2003), and whether people's behaviour maps on to these predictions. This exploration would require the integration of both Assertion and Expertise variables, factors which are severely underexplored in the empirical literature (Hahn, Harris, & Corner, 2016).

4.3 Experiment 4.2: Priors, Assertion, and the probability of the antecedent

Experiment 4.2 investigated the intuition that, when we hear the assertion of a conditional, we may learn something about the antecedent depending on our prior beliefs.

4.3.1 Methods

Participants. 354 participants (161 female, 1 non-binary; average age 35.38 years) completed the task; 9 participants had previously been excluded since their first language was not English. Participants were recruited via the intermediary MTurk Data (www.mturkdata.com). High qualifications were set for the task to improve the quality of the data and maximize the number of native English speakers: participants had to be resident in the US, and have an overall approval rating of 99% for 1,000 previously completed tasks. Participants received a small fee, chosen to exceed the US minimum wage per minute.

Materials. The materials were adapted from the previous studies. The materials manipulated two variables: Assertion (Null, Single, Multiple); Prior (Low, High). The prior probability of the antecedent was manipulated by contexts. The full items are as follows. Note that the brackets contain the manipulation for multiple sources and for high prior.

- (1) Adam (...Barbara, Nick, and Sue) is (are) at a large car dealership which specializes in mid-range cars. He (They) tells (tell) you, 'If a car is a Rolls Royce (Chevrolet), then it's black.' What's the probability that a car in the dealership is a Rolls Royce (Chevrolet)?
- (2) Jo (... Aaron, Zoe and Felix) is/are at a medical clinic. She (They) tells (tell) you, 'If a patient has bubonic plague (a common cold),

RATIONALITY, PRAGMATICS, AND SOURCES

- then they'll make a good recovery.' What's the probability that a patient has bubonic plague (a common cold)?
- (3) Mike (...Isabel, Elliot and Amy) is (are) at a steakhouse. He (They) tells (tell) you, 'If a customer is eating tofu (steak), then it's gluten free.' What's the probability that a customer is eating tofu (steak)? Please make your response on a scale from 0 to 10, where 0 is not at all possible and 10 is certain.
- (4) Anna (...Owen, Julia and Nathan) is (are) at a zoo. She (They) tells (tell) you, 'If the zoo has black rhinoceroses (seals), then they're by the main building.' What's the probability that the zoo has black rhinoceroses (seals)?
- (5) Colin (...Tom, Sarah and Leo) is (are) visiting a Liberal Arts College. He (They) tells (tell) you, 'If Lisa, a student, is majoring in astrophysics (an arts subject), then she's working late in the library.' What's the probability that Lisa is majoring in astrophysics (an arts subject)?
- (6) Lizzy (...Holly, Jamie and Maria) is (are) at a veterinary clinic. She (They) tells (tell) you, 'If Susie is an African grey parrot (dog), then she's had her operation.' What's the probability that Susie is an African grey parrot (dog)?
- (7) Lorraine (...Eric, Robby and Emma) is (are) at a coffee shop. She (They) tells (tell) you, 'If Dan, another customer, is drinking a decaffeinated (caffeinated) coffee, then he has work to do.' What's the probability that Dan is drinking a decaffeinated (caffeinated) coffee?

As in previous studies, in the null condition, the first sentence was replaced by ‘Imagine you are at...’, followed directly by the probability question.

Procedure. This task focused on the probability of the antecedent; hence there was a single dependent measure, $P(\text{Antecedent})$, measured on a scale from 0 to 10. After giving informed consent, participants were assigned, in a round-robin fashion, to a condition. Participants read the same instructions as for Experiments 2.1 to 2.4. Participants rated all 7 items, which appeared in a random order.

4.3.2 Results & Discussion.

The key prediction was for an interaction between Prior and Assertion.

Figure 4.2 shows the descriptive data for the effects of Assertion and Prior on the probability of the antecedent.

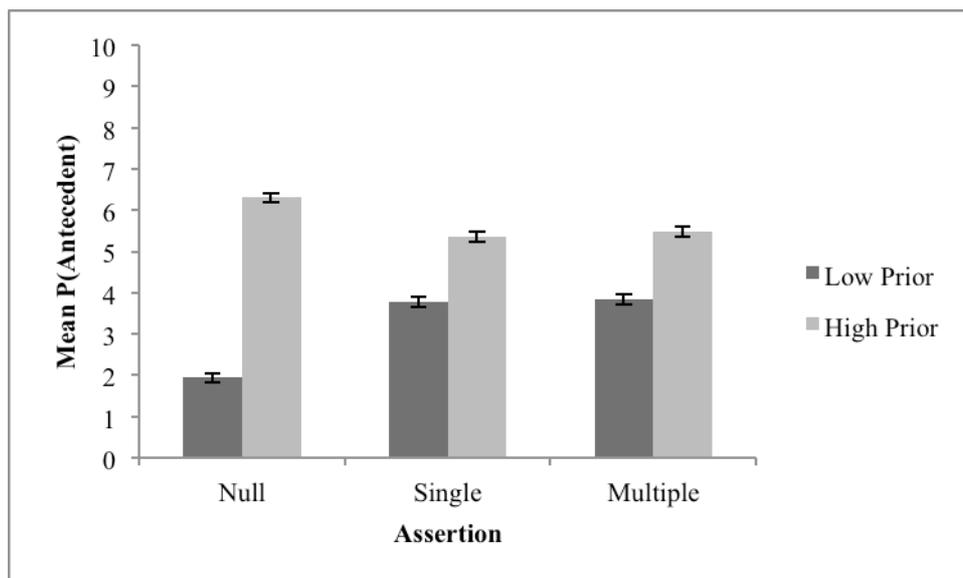


Figure 4.2. Effect of Assertion and Prior on $P(\text{Antecedent})$; error bars are standard error. The descriptive data suggest an interaction. With low priors, there appears to be a linearly increasing trend, from the null condition ($M = 1.95$) to the single condition ($M = 3.77$) and the multiple condition ($M = 3.83$). With high priors, in contrast, there appears to be a decrease from the null condition ($M = 6.31$) to the single condition ($M = 5.37$) and then a slight increase to the multiple condition ($M = 5.48$). This

picture is confirmed by mixed-effects modelling; these models include the random intercepts only for topics and items, because of convergence problems. Including the interaction term significantly improved fit over the null model, $\chi^2(2) = 81.12, p < .001$. Table 4.3 reports the estimates of the fixed effects.

Table 4.3. Fixed effects of Prior, Assertion, and Interaction

Fixed Effect	Parameter	95% CI
Intercept	$b = 1.95$	1.27, 2.63
Prior (High)	$b = 4.36$	3.90, 4.81
Assertion (Single)	$b = 1.82$	1.36, 2.28
Assertion (Multiple)	$b = 1.88$	1.42, 2.33
Prior(H)*Assertion(S)	$b = -2.76$	-3.40, -2.11
Prior(H) * Assertion(M)	$b = -2.70$	-3.35, -2.05

To explore the interaction further, the data were split for low and high priors, and separate analyses were run to explore the effect of Assertion. The models, from here on, included the full random-effects structure. For the low prior data, including Assertion significantly improved fit over the null model, $\chi^2(2) = 18.04, p < .001$. Table 4.4 reports the fixed effects of Assertion for the low prior data.

Table 4.4. Fixed effects of Assertion at low prior

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 1.95$	1.13, 2.77
Single	$b = 1.82$	1.24, 2.40
Multiple	$b = 1.88$	1.29, 2.46

Pairwise comparisons were performed on the least-squares means, using the Tukey correction for multiple comparisons. The increase from null to single conditions was significant, $M = 1.82, 95\% \text{ CI } [1.06, 2.57], t(25.39) = 5.99, p < .001$. So too was the increase from null to multiple conditions, $M = 1.88, 95\% \text{ CI } [1.13, 2.62], t(40.71) = 6.13, p < .001$. The increase from single to multiple conditions was not, however, significant, $M = .06, 95\% \text{ CI } [-.6, .72], t(55.24) = .21, p = .98$.

RATIONALITY, PRAGMATICS, AND SOURCES

For the high prior data, including Assertion also significantly improved fit over the null model, $\chi^2(2) = 8.47, p = .01$. Table 4.5 reports the estimates of the fixed effects of Assertion for the high prior data.

Table 4.5. Fixed effects of Assertion at high prior

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 6.31$	5.35, 7.26
Single	$b = -.94$	-1.72, -.16
Multiple	$b = -.82$	-1.32, -.33

As above, pairwise comparisons were performed on the least-squares means, using the Tukey correction for multiple comparisons. The decrease from null to single conditions was not significant, $M = -.94$, 95% CI [-2.03, .16], $t(14.80) = -2.23, p = .10$. The decrease from null to multiple conditions, in contrast, was significant, $M = -.82$, 95% CI [-1.46, -.19], $t(32.39) = -3.20, p = .008$. The difference between single and multiple conditions was not significant, $M = .11$, 95% CI [-.65, .88], $t(25.11) = .37, p = .93$.

Although it could be argued that a fuller picture would be given by investigating the effect of Prior at each level of Assertion, such analyses are not relevant for our predictions, which concerned only the effect of Assertion. We therefore omit the analyses here. The same applies to all the experiments below.

The data vindicate the intuitions reported earlier in this chapter: the effect of Assertion depends on the prior probability of the antecedent. With low priors, there is a reliable increase in the ratings of the probability of the antecedent when a conditional is asserted. With high priors, there is a reliable decrease.

4.4 Experiment 4.3: Priors, Expertise, and the probability of the antecedent

4.4.1 Method

Participants. 357 participants (162 female, 1 non-binary; average age 34.56 years) completed the task; 9 participants had previously been excluded since their first language was not English. Participants were recruited in the same manner as above, and remunerated in the same way.

Materials. The materials manipulated Prior (Low, High) and Expertise (Null, Inexpert, Expert). The same contexts and items were used as for Experiment 4.2. The Expertise manipulation was implemented in the following way, replacing the text for the Assertion manipulation.

- (1) A customer/manager tells you...
- (2) A medical student/professor of medicine tells you...
- (3) The regular handyman/the head chef tells you...
- (4) A deliveryman/a veterinary nurse tells you...
- (5) A five-year-old visitor/zoo keeper tells you...
- (6) The canteen manager/personal tutor of Lissa, a student, tells you...
- (7) An old friend of Dan, another customer, who had not seen Dan for 10 years/the girlfriend of Dan, another customer, tells you...

The null condition was implemented in the same was as in Experiment 4.3.

Procedure. The procedure was the same as for Experiment 4.3.

4.4.2 Results & Discussion

The key prediction was an interaction between Prior and Expertise. Figure 4.3 shows the descriptive data for the effect of Prior and Expertise on the probability of the antecedent.

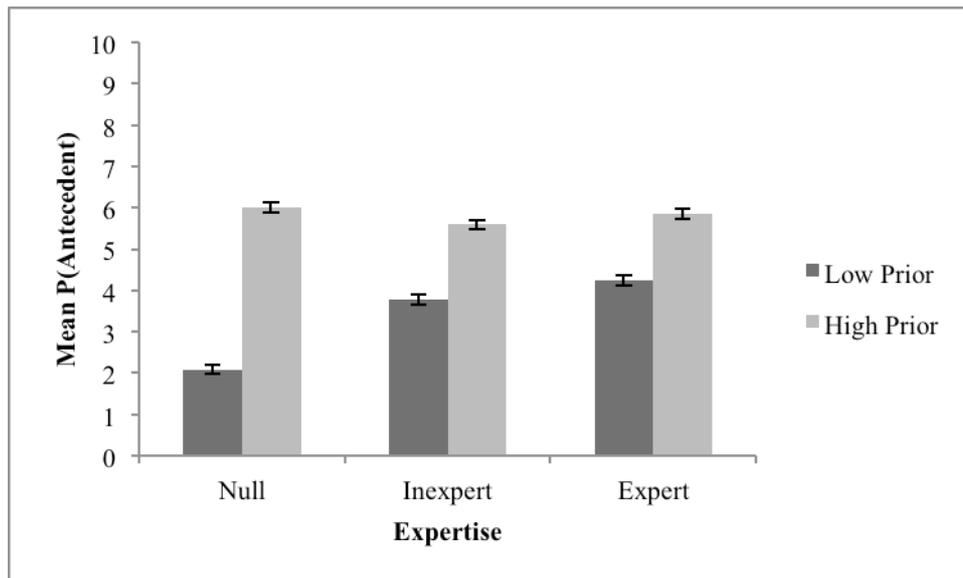


Figure 4.3. Effect of Prior and Expertise on P(Antecedent); error bars are standard error. The descriptive data suggest an interaction. As with the Assertion data, when the prior is low, there appears to be a linear trend of Expertise, from the null condition ($M = 2.08$) to the inexpert condition ($M = 3.78$) and the expert condition ($M = 4.25$). When the prior is high, there is a decrease from the null condition ($M = 6$) to the inexpert condition ($M = 5.6$) and a slight increase to the expert condition ($M = 5.86$). Mixed-effects modelling supported this picture. The models, here, included the full random-effects structure. Including the interaction term significantly improved fit over the null model, $\chi^2(2) = 10.11, p = .006$.

Table 4.6 reports the estimates of the fixed effects.

Table 4.6. Fixed effects of Prior, Expertise, and Interaction on P(Antecedent)

Fixed Effect	Parameter	95% CI
Intercept	$b = 2.08$	1.16, 2.99
Prior (High)	$b = 3.92$	2.92, 4.93
Expertise (Inexpert)	$b = 1.70$	1.00, 2.40
Expertise (Expert)	$b = 2.17$	1.40, 2.94
Prior(H)*Expertise(I)	$b = -2.10$	-3.13, -1.08
Prior(H) * Expertise(E)	$b = -2.31$	-3.39, -1.23

RATIONALITY, PRAGMATICS, AND SOURCES

To explore the interaction further, the data were split for low and high priors, and separate analyses were run to explore the effect of Expertise. For the low prior data, including Expertise significantly improved fit over the null model, $\chi^2(2) = 13.13$, $p = .001$. Table 4.7 reports the estimates of the fixed effects.

Table 4.7. Fixed effects of Expertise on P(Antecedent) at low prior

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 2.08$	1.16, 3.00
Inexpert	$b = 1.70$.99, 2.42
Expert	$b = 2.17$	1.39, 2.95

Pairwise comparisons were conducted on the least-squares means, using the Tukey correction for multiple comparisons. The increase from null to inexpert conditions was significant, $M = 1.70$, 95% CI [.74, 2.67], $t(19.29) = 4.48$, $p < .001$. The increase from null to expert conditions was also significant, $M = 2.17$, 95% CI [1.12, 3.23], $t(19.72) = 5.22$, $p < .001$. The increase from inexpert to expert conditions was not significant, $M = .47$, 95% CI [-.18, 1.11], $t(44.72) = 1.76$, $p = .19$.

For the high prior data, including Expertise did not significantly improve fit over the null model, $\chi^2(2) = 1.90$, $p = .39$. Table 4.8 reports the estimates of the fixed effects.

Table 4.8. Fixed effects of Expertise on P(Antecedent) at high prior

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 6.00$	5.00, 7.00
Inexpert	$b = -.40$	-1.15, .35
Expert	$b = -.14$	-.92, .64

These data generally support the intuition that the effect of Expertise depends on the prior probability of the antecedent. The data are similar to the Assertion data: as above, with low priors there is a reliable increase in the probability of the

antecedent; with high priors, there is a decrease but, unlike above, the effect is not significant.

4.5 Experiment 4.4: Priors, Assertion, and the probability of the consequent

4.5.1 Methods

Participants. 367 participants (163 female, 4 other; average age 33.98 years) completed the task; 5 participants had previously been excluded since their first language was not English. Participants were recruited and remunerated in the same way as in Experiments 4.1 to 4.3

Materials. The items, here, were adapted from Experiment 4.3. The antecedent and consequents of the conditional were simple swapped. One item had to be adapted. The previous experiment contained the following item:

Jo (... , Aaron, Zoe and Felix) is/are at a medical clinic. She (They) tells (tell) you, 'If a patient has bubonic plague (a common cold), then they'll make a good (poor) recovery.' What's the probability that a patient has bubonic plague (a common cold)?

This item does not work when the antecedent and consequent are swapped. It was therefore replaced with the following item. In brackets is the text for the multiple assertion condition and for the low prior condition.

Jo (... Aaron, Zoe and Felix) is (are) at a medical clinic with an excellent recovery rate. She (They) tells (tell) you, 'If a patient on the ward, has malaria, then he will make a good recovery.' What's the probability that a patient will make a good (poor) recovery?

Procedure. The design manipulated Assertion (Null, Single Assertion, Multiple Assertion) and Prior Probability (Low, High). This task focused on the probability of the consequent; hence there was a single dependent measure,

P(Consequent), measured on a scale from 0 to 10. The task was between-subjects.

After giving informed consent, participants received the same instructions as in

Experiment 4.2, and were assigned, in a round-robin fashion, to a condition.

Participants saw all seven items, with the ordering being randomized.

4.5.2 Results & Discussion

The key prediction was an interaction between Prior and Assertion. Figure 4.4 shows the descriptive data for the effects of Prior and Assertion on the probability of the consequent.

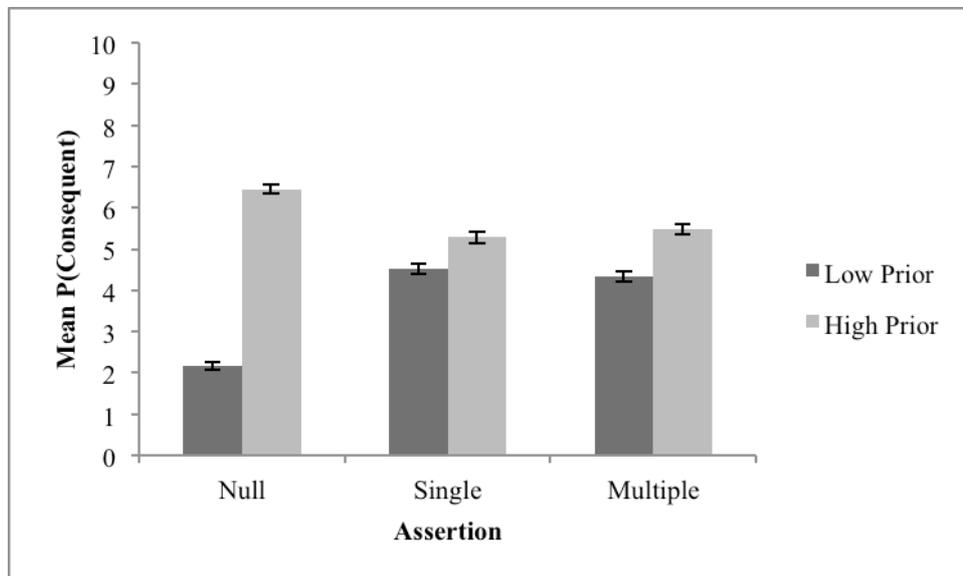


Figure 4.4. Effects of Prior and Assertion on P(Consequent); error bars are standard error

These data suggest an interaction. With low priors, there is a trend for Assertion to increase ratings of P(Consequent): an increase from the null condition ($M = 2.15$) to the single condition ($M = 4.51$), which tails off somewhat with the multiple condition ($M = 4.34$). With high priors, there is the opposite trend: a decrease from the null condition ($M = 6.45$) to the single condition ($M = 5.29$), which tails off with the multiple condition ($M = 5.48$).

A mixed-effects model (with full random-effects structure) showed that including the interaction term produced a significant improvement in fit over the null

RATIONALITY, PRAGMATICS, AND SOURCES

model (i.e. the model without the interaction term), $\chi^2(2) = 120.16, p < .001$. Table 4.9 shows the estimates of the fixed effects.

Table 4.9. Fixed effects of Prior, Assertion, and Interaction on P(Consequent)

Fixed Effect	Parameter	95% CI
Intercept	$b = 2.15$	1.37, 2.93
Prior (High)	$b = 4.30$	3.53, 5.07
Assertion (Single)	$b = 2.36$	1.71, 3.00
Assertion (Multiple)	$b = 2.19$	1.44, 2.93
Prior(H)*Assertion(S)	$b = -3.53$	-4.17, -2.88
Prior(H) * Assertion(M)	$b = -3.16$	-3.79, -2.52

To explore the interaction further, the data were split for low and high priors, and separate analyses were run to explore the effect of Assertion. For the low prior data, including Assertion significantly improved fit over the null model, $\chi^2(2) = 18.19, p < .001$. Table 4.10 reports the estimates of the fixed effects.

Table 4.10. Fixed effects of Assertion on P(Consequent) at low prior

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 2.15$	1.26, 3.05
Single	$b = 2.36$	1.73, 2.99
Multiple	$b = 2.38$	1.61, 3.15

Pairwise comparisons were performed on the least-squares means, using the Tukey correction for multiple comparisons. The increase from null to single conditions was significant, $M = 2.36$, 95% CI [1.53, 3.19], $t(27.87) = 7.05, p < .001$. The increase from null to multiple conditions was also significant, $M = 2.38$, 95% CI [1.32, 3.45], $t(15.85) = 5.78, p < .001$. The increase from single to multiple conditions was not significant, $M = .02$, 95% CI [-.60, .64], $t(41.48) = .10, p = .99$.

For the high prior data, including Assertion produced a marginally significant improvement in fit over the null model, $\chi^2(2) = 5.48, p = .06$. Table 4.11 reports the estimates of the fixed effects.

Table 4.11. Fixed effects of Assertion on P(Consequent) at high prior

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 6.45$	5.35, 7.55
Single	$b = -1.17$	-2.01, -.32
Multiple	$b = -1.12$	-2.06, -.17

Since the improvement was so close to significance, pairwise comparisons were performed on the least-squares means, using the Tukey correction for multiple comparisons. The decrease from null to single conditions was significant, $M = -1.17$, 95% CI [-2.32, -.01], $t(18.14) = -2.58$, $p = .05$. The decrease from null to multiple conditions was not significant, $M = -1.12$, 95% CI [-2.45, .21], $t(14.21) = -2.19$, $p = .11$. The decrease from single to multiple conditions was also not significant, $M = .05$, 95% CI [-.58, .68], $t(41.20) = .19$, $p = .98$.

This experiment shows that the effect of Assertion on the probability of the consequent depends on the prior probability. The effect is equivalent to the effect found for the probability of the antecedent, though the decreasing trend with high priors is less statistically reliable.

4.6 Experiment 4.5: Priors, Expertise, and the probability of the consequent

4.6.1 Methods

Participants. 356 participants (173 female; average age 35.06 years) completed the task; 9 participants had previously been excluded since their first language was not English. This experiment used the same system as above for recruiting and remunerating participants.

Materials. These were the conditions for Experiment 4.4, using the Expertise manipulation from Experiment 4.3.

Procedure. The design manipulated Assertion (Null, Inexpert Source, Expert Source) and Prior Probability (Low, High). This task focused on the probability of

the consequent; hence there was a single dependent measure, $P(\text{Consequent})$, measured on a scale from 0 to 10. The task was between-subjects. After giving informed consent, participants read the same instructions as for Experiments 4.2, 4.3, and 4.4, and were assigned, in a round-robin fashion, to a condition. Participants saw all seven items in a randomized order.

4.6.2 Results & Discussion

The key prediction was an interaction between Prior and Assertion. Figure 4.5 shows the descriptive data for the effects of Prior and Assertion on the probability of the consequent.

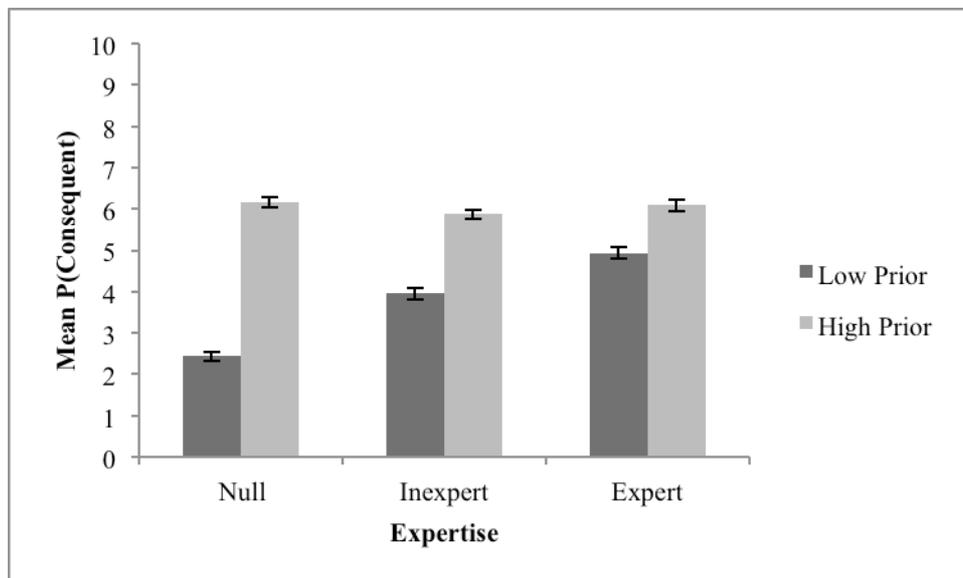


Figure 4.5. Effects of Prior and Expertise on $P(\text{Consequent})$; error bars are standard error

The data against suggest an interaction. With low priors, there is a trend for Expertise to increase ratings of $P(\text{Consequent})$, from the null condition ($M = 2.44$) to the inexpert condition ($M = 3.96$) and the expert condition ($M = 4.93$). With high priors, in contrast, the data seem rather flat for the null ($M = 6.17$), inexpert ($M = 5.87$) and expert ($M = 6.09$) conditions. A mixed-effects model, run using the full random-effects structure, showed that including the interaction term significantly

RATIONALITY, PRAGMATICS, AND SOURCES

improved fit, $\chi^2(2) = 48.66$, $p < .001$. Table 4.12 reports the estimates of the fixed effects.

Table 4.12. Fixed effects of Prior, Expertise, and Interaction on P(Consequent)

Fixed Effect	Parameter	95% CI
Intercept	$b = 2.44$	1.63, 3.25
Prior (High)	$b = 3.74$	3.03, 4.44
Expertise (Inexpert)	$b = 1.52$.86, 2.18
Expertise (Expert)	$b = 2.49$	1.71, 3.26
Prior(H)*Expertise(I)	$b = -1.82$	-2.53, -1.11
Prior(H) * Expertise(E)	$b = -2.56$	-3.28, -1.85

To explore the interaction further, the data were split for low and high priors, and separate analyses were run to explore the effect of Expertise. For the low prior data, including Expertise significantly improved fit over the null model, $\chi^2(2) = 20.79$, $p < .001$. Table 4.13 reports the estimates of the fixed effects.

Table 4.13. Fixed effects of Expertise on P(Consequent) at low prior

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 2.44$	1.58, 3.29
Inexpert	$b = 1.52$.85, 2.18
Expert	$b = 2.45$	1.86, 3.12

Pairwise comparisons were performed on the least-squares means, using the Tukey correction for multiple comparisons. The increase from null to inexpert conditions was significant, $M = 1.52$, 95% CI [.66, 2.38], $t(33.39) = 4.33$, $p < .001$. The increase from null to expert conditions was also significant, $M = 2.49$, 95% CI [1.68, 3.30], $t(35.64) = 7.51$, $p < .001$. Lastly, the increase from inexpert to expert conditions was significant, $M = .97$, 95% CI [.31, 1.63], $t(41.67) = 3.57$, $p = .003$.

With high priors, including Expertise did not significantly improve model fit, $\chi^2(2) = 1.12$, $p = .57$. Table 4.14 reports the estimates of the fixed effects.

Table 4.14. Fixed effects of Expertise on P(Consequent) at high prior

Fixed Effect	Parameter	95% CI
Intercept (Null)	$b = 6.17$	4.87, 7.47
Inexpert	$b = -.30$	-1.29, .68
Expert	$b = -.07$	-1.20, 1.06

The data for this experiment are consistent with those for the probability of the antecedent: as above, with low priors, there is a reliable increase in the probability of the antecedent; with high priors, there is decrease but this is not reliable.

4.7 General Discussion

The data above round off this thesis' experimental work on testimonial conditionals; later chapters will consider related phenomena. We turn, now, to a summary of these three chapters' data and to their philosophical and psychological implications.

It will prove useful, here, to recapitulate the main findings. Chapters 2 and 3 offered consistent evidence that both assertion and source expertise reliably increased the conditional probability, $P(\text{Consequent}|\text{Antecedent})$. With these experimental materials, assertion and expertise left the probability of the antecedent and the probability of the consequent relatively untouched. Chapter 3 also demonstrated that both assertion and expertise decreased interval estimates of the conditional probability: that is, participants gave narrower, more precise estimates. Although the assertion experiments generally showed a plateauing effect, whereby there was no reliable increase from single to multiple assertion, Experiment 4.1 suggested that this plateauing may depend on the ambiguity of the experimental items in that multiple-assertion condition. Lastly Experiments 4.2 to 4.5 suggested that assertion and expertise can shift both the probability of the antecedent and the

RATIONALITY, PRAGMATICS, AND SOURCES

probability of the consequent in appropriate contexts. Assertion and expertise reliably increased the probability of the antecedent and the probability of the consequent when their respective prior probabilities were low. Assertion reliably decreased the probability of the antecedent, and marginally so the probability of the consequent, when their respective prior probabilities were high. Expertise did not have a reliable effect when the prior probabilities were high. A more reliable effect may result from tasks whose materials succeed in generating higher priors: ratings in the null condition reached only around 6 on a scale from 0 to 10.

Chapters 2 and 3 suggested that their data are consistent with the leading psychology theory of the conditional, the suppositional theory. This suggestion can now be tested more rigorously against the full set of data. The suppositional theorist can straightforwardly account for the data in Chapter 2. On their theory, participants supposed that the antecedent was true, and judged the conditional probability. Participants took the assertion of the conditional to mean that the conditional probability was high. They may also have drawn on the number of people making the assertion (see Experiment 4.1). When participants had information about the source's expertise available to them, they drew on this to judge the conditional probability.

Other data are less easy to explain with the suppositional theory. Take, first, the data in Chapter 3. The data on slider averages pose no problem for the suppositional theorist, because they are amenable to the same explanation as the point-values data in Chapter 2. Matters are less straightforward for the slider ranges: these data do not contradict anything in the suppositional theory, but nor are they predicted by anything in the theory. It remains to be seen whether the suppositional theory should simply be extended to accommodate interval estimates and whether, if

so, this extension has any profound effect, positive or otherwise, on the theory's predictions.

The data from experiments 4.2 to 4.5 also do not find an obvious explanation in the suppositional theory. These data suggest that, when someone hears a conditional, they may revise more than just the conditional probability: that the probability of the antecedent and the probability of the consequent can also shift³¹. The crucial factor is the relevant prior probability. Again, these data do not contradict anything in the suppositional theory. To illustrate this point, consider the ratio definition of conditional probability as applied to the conditional 'If a car is a Mercedes, then it's black'. Assume that the conditional refers to a large car showroom with cars of different makes and different colours, and that the reference class is the cars in this showroom:

$$P(\text{Black} \mid \text{Mercedes}) = \frac{P(\text{Mercedes} \ \& \ \text{Black})}{P(\text{Mercedes})}$$

The conditional probability, in other words, is the proportion of Mercedes cars in the showroom that are black. Given the data in Chapters 2 to 4, we are interested only in cases in which the conditional probability rises. The conditional probability rises if the proportion of Mercedes cars which are black rises. But this change in the proportion is compatible with various actions by the management. Take some of the simplest changes. The number of Mercedes and the number of black cars can stay the same: the management can re-spray some Mercedes black and the same number

³¹ Experiments 4.2 to 4.5 were designed to test whether $P(\text{Antecedent})$ and $P(\text{Consequent})$ could be made to shift. They were not designed to study exhaustively the effects of priors. So, it remains to be seen whether the priors for one probability affect the other probabilities. Plausibly they do. For instance, a conditional with low $P_0(\text{Consequent})$ could be taken as a kind of *reductio ad absurdum*. Consequently, $P(\text{Antecedent})$ might decrease.

of black non-Mercedes another colour. The number of Mercedes can increase - the management can keep all existing Mercedes and buy in some new black ones – or decrease – the management can keep all its black Mercedes and sell some non-black ones. The number of black cars can vary freely as long as the black Mercedes are untouched. But, of course, if all the cars are black, then the conditional probability will be one. As this scenario illustrates, the problem is not that the suppositional theory is incompatible with change in $P(\text{Antecedent})$ and $P(\text{Consequent})$. It is, rather, that there is nothing in the theory to explain the systematicity of the changes.

4.7.1 Modelling testimonial conditionals

There may, yet, be a fairly straightforward probabilistic account for both the increase in $P(\text{Consequent}|\text{Antecedent})$ and the systematic changes in $P(\text{Antecedent})$ and $P(\text{Consequent})$. This account employs formal Bayesian models of testimony which were alluded to in Chapter 2. A simple model will serve as the entry point for this discussion. Figure 4.6 shows a simple model of testimony from Bovens and Hartmann (2003).

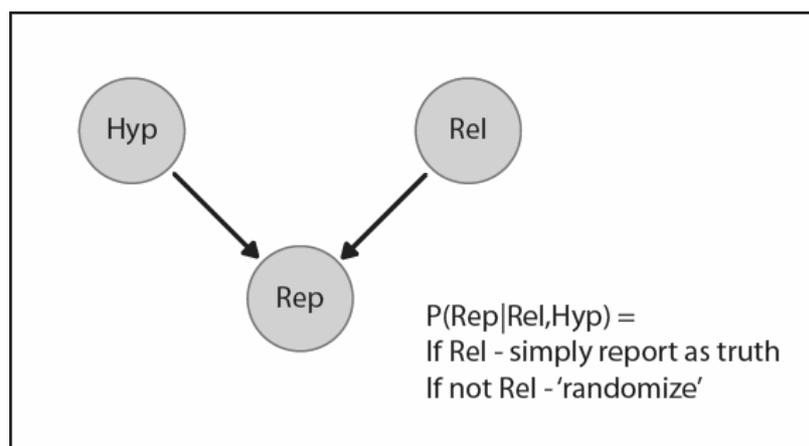


Figure 4.6. Bayesian belief network from Bovens and Hartmann (2003)

This is a simple Bayesian belief network: the nodes represent random variables; the arcs represent probabilistic relationships between variables (for

RATIONALITY, PRAGMATICS, AND SOURCES

discussion, see, e.g., Korb & Nicholson, 2011). The variable *Hyp*, here, is a proposition about the world, and has the states true and false. The variable *Rel*, here, is the reliability of the speaker: it, too, has the states true (reliable) and false (unreliable). Lastly, the variable *Rep* is the report that ‘*Hyp* is true’: it has the states true (the report is made) and false (the report is not made).

To illustrate this simple model, consider the different positions of a source and hearer, assuming that the proposition under discussion is true. Take, first, the position of the source. Does the source report that ‘*Hyp* = True’? This will depend on two factors: the source’s beliefs about the state of the world, and the source’s reliability. In this situation, these factors have an independent influence on the node *Rep*. The source’s beliefs about the world are captured by a distribution over the binary states of *Hyp*. Similarly, the source’s reliability is represented by a distribution over the binary states of *Rel*. On this simple model, a reliable source will simply report ‘*Hyp*=True’. An unreliable source, in contrast, will randomize: s/he will, in effect, toss a coin to decide whether to report that ‘*Hyp* is true’. This relationship is defined in the conditional probability table for the node *Rep*, which is shown in Table 4.15.

Table 4.15. Conditional probability table for Bovens Hartmann network

Hyp	Reliable		Unreliable	
	True	False	True	False
Rep = True	1	0	.5	.5
Rep = False	0	1	.5	.5

This network accommodates simple cases, where $P(Hyp=True)$ and $P(Rel=True)$ are both only 0 or 1, and subtler cases, where $P(Hyp=True)$ and $P(Rel=True)$ range anywhere between 0 and 1.

Now take the position of the hearer. Two states of the network are relevant here: before and after a source's testimony. In the initial state, the hearer has prior beliefs about the world represented by a distribution over the binary states of *Hyp*, and an initial perception of the source's reliability represented by a distribution over the binary states of *Rel*. These variables are marginally independent. However, when the source reports that '*Hyp*=True', the hearer intervenes in the network and sets the state of *Rep* to true. Thereafter, the variables *Hyp* and *Rel* are conditionally dependent. The hearer revises his/her distributions for both variables. Information from the *Hyp* variable flows through to the *Rel* variable. If the hearer's initial belief was $P(Hyp=True) > .5$, then the probability of $P(Rel=True)$ will increase; but if the hearer's initial belief was $P(Hyp=True) < .5$, then the probability of $P(Rel=True)$ will decrease. Conversely, information from the *Rel* variable flows through to the *Hyp* variable. If the hearer's initial reliability judgment was $P(Rel=True) = 0$, then the source will have no effect on the belief in $P(Hyp=True)$. If the hearer's initial reliability judgment was $P(Rel=True) > 0$, then $P(Hyp=True)$ will increase.

This simple model can prove a valuable source of hypotheses for empirical research (see Chapter 6). However, it has little, if any, application to testimonial conditionals, for three principal reasons. While it might be tempting to simply assign the *Hyp* node to the conditional, this assignment would ignore the philosophical controversy over whether conditionals are propositions (for discussion, see Douven, 2012). Relatedly, the assignment would obscure the fact that conditionals refer to a (potentially probabilistic) relationship between two variables: the antecedent and consequent. Finally, the assignment would run counter to the data discussed in this chapter: since the model would not represent antecedent and consequent separately, it could not represent changes to their probabilities either.

RATIONALITY, PRAGMATICS, AND SOURCES

A more sophisticated approach has been proposed by Stephan Hartmann (reported in Collins, Krzyżanowska, Hartmann, Wheeler, & Hahn, 2017). The approach uses a baseline model, shown Figure 4.7.

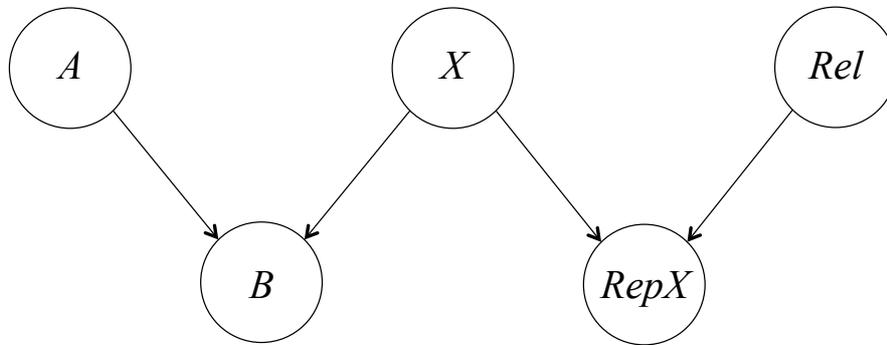


Figure 4.7. Baseline Bayesian belief network for experimental data

The variable *A* is the antecedent of a conditional, and has the values ‘The antecedent occurs’ and ‘The antecedent does not occur’. The variable *B* is the consequent, and has the values ‘The consequent occurs’ and ‘The consequent does not occur’. The variable *X* is the conditional, and has the values ‘The indicative conditional “If *A*, *B*” holds’ and ‘The indicative conditional “If *A*, *B*” does not hold’. Including such a variable is a strategy which originates in solutions to the Old Evidence Problem in epistemology (these will be discussed further below). Importantly, the variable is neutral with respect to theories of the conditional: the model, therefore, makes no major theoretical assumptions. The variable *Rel* is the reliability of the source, and has the values ‘The source is reliable’ and ‘The source is not reliable’. Lastly, the variable *RepX* is the report of the conditional, and has the values ‘The source reports *X*’ and ‘The source does not report *X*’. The arcs, in the model, once again represent probabilistic relationships between the variables.

RATIONALITY, PRAGMATICS, AND SOURCES

The network also requires the probability distribution to be specified. Nodes A , X , and Rel are root nodes; as such, they can be assigned prior probabilities. There are, in addition, two conditional probability tables. Firstly, for node B ³²:

Table 4.16. Conditional probability table for node B

Conditional (X)	Holds		Not	
	Occurs	Not	Occurs	Not
Consequent (B)				
Occurs	1	β	α	γ
Not	0	$1 - \beta$	$1 - \alpha$	$1 - \gamma$

Here, no particular values are assumed for most parameters; however, it is assumed that X respects *modus ponens*. This assumption can be relaxed to allow for uncertainty, but I will not discuss this relaxation in this thesis. The second conditional probability table is for node $Rep X$.

Table 4.17. Conditional probability table for node $Rep X$

Conditional (X)	Holds		Not	
	Reliable	Not	Reliable	Not
Report X	1	μ	0	μ
Not Report X	0	$1 - \mu$	1	$1 - \mu$

Here it is assumed that a reliable source will always report that the conditional holds when it does and never when it does not. In contrast, an unreliable source randomizes with probability μ . Note the similarity with Table 4.15 when μ is .5.

³² In other words, $P(B|A, X) = 1$, $P(B|A, \sim X) = \alpha$, $P(B|\sim A, X) = \beta$, and $P(B|\sim A, \sim X) = \gamma$. And likewise, $P(RepX|X, Rel) = 1$, $P(RepX, X, \sim Rel) = \mu$, $P(RepX, \sim X, Rel) = 0$, and $P(RepX|\sim X, \sim Rel) = \mu$.

How does this model bear on the experimental data described above? This model was developed as a baseline model to account for the data in Experiments 2.1 to 2.4. Consider, first, node A , the antecedent. A is marginally independent of $RepX$: A 's influence is blocked by B . Thus, A 's probability distribution will affect only the probability distribution over the states of B , unless certain interventions are made in the network. Imagine, now, that a source asserts a conditional: we intervene so that $P(Rep X = \text{'Reports } X\text{'}) = 1$. This intervention makes nodes X and Rel conditionally dependent. Consequently, the distributions change for nodes B , X , and Rel . However, the distribution for A remains untouched. When a conditional is asserted, the probability of the antecedent remains the same but the conditional probability increases.

A second crucial aspect of the data was that the probability of the consequent did not shift. At first blush, this point is problematic for the network above because, as we have seen, the probability of the consequent – that is, the distribution over the states of node B – routinely shifts on the assertion of the conditional (intervention at node $RepX$). However, the probability of the consequent will not shift as long as two conditions hold: (1) that $\beta \approx \gamma$ (i.e. $P(B|\neg A, X) \approx P(B|\neg A, \neg X)$); and (2) that $\alpha \approx 1$ (i.e. $P(B|A, \neg X) \approx 1$). To elaborate, condition (1) requires that the following probabilities are approximately the same: the probability of the consequent given that the antecedent does not hold and the indicative conditional does; and the probability of the consequent given that the antecedent does not hold and the indicative conditional does not hold. Condition (2) requires that the following probability is approximately

RATIONALITY, PRAGMATICS, AND SOURCES

1: the probability of the consequent given that the antecedent holds and the indicative conditional does not³³.

How reasonable is it to assume that the conditions hold for the data in Chapter 2? These assumptions have not been empirically tested, but their plausibility can nevertheless be assessed. Consider, first, β : the probability of the consequent given that the antecedent does not hold and the indicative conditional does. Plausibly, participants were non-committal on this probability: this probability occupies one of the ‘irrelevant’ or ‘void’ cells in the defective truth table, which, as we saw in Chapter 2, people arguably endorse. Moreover, participants would have no real information to judge how likely the consequent was in the relevant situation, since the scenarios were fictional. Consider, now, γ : the probability of the consequent given that the antecedent does not hold and the indicative conditional does not hold. Again, for the present studies, participants would likely be non-committal on this probability: they had no information to judge the probability of the consequent in this situation either. It is possible, then, that participants treated the probabilities as

³³ These conditions were explored in simulations in Hugin using priors for the *A*, *X*, and *Rel* that roughly corresponded to those in the data (including related data sets which are not reported in this thesis). The values of β and γ were initially set to .5. Even small departures from the conditions resulted in change to the probability of the consequent. When condition (1) held, but assumption (2) did not, then the probability of the consequent increased (though small changes to α led to small, tolerable increases in the probability of the consequent). When condition (2) held, but assumption (1) did not the probability of the consequent increased when β was greater than γ , and decreased when γ was greater than β .

approximately the same. It is not clear, however, whether these assumptions would have held if participants could have deployed their real-world knowledge.

We turn, finally, to, α : the probability of the consequent given that the antecedent holds and the indicative conditional does not. This probability judgement requires us to separate judgement of the conditional probability and judgement of whether the indicative conditional holds. But it has been stressed, in previous chapters, how close these judgements seem to be. Nevertheless, separation is needed to avoid damaging self-reference in the model: the node cannot simply refer to the arc between nodes A and B (for discussion, see Collins et al., 2017). To make sense, node X must refer to other (additional, new) information not presently summarized by the arc between A and B . How plausible is inclusion of this node? We can answer this question by considering the Old Evidence Problem, which was raised by Glymour (1980). The problem is as follows. When proposing a new theory, we will likely consider it a virtue for the theory to explain existing evidence. But this virtue raises a problem for a Bayesian epistemology. Bayesian epistemology takes as a measure of support $P(H|E) > P(H)$. But if the evidence is already known then $P(H|E) = P(H)$ (for discussion, see Howson, 1991) and, in Bayesian terms, the evidence cannot be said to support the theory. Garber (1983) proposed adding a variable in a similar way to Stephan Hartmann's model (see also Hartmann and Fitelson, 2015). The extra variable represents the discovery of new information: the deductive inference (or prediction) from the theory to the data. Similarly, in the present case, the extra variable could represent the discovery of new information: that the conditional holds. Since the model is neutral on the meaning of the conditional, though, it is also neutral on the nature of this new information.

The model in Figure 4.7 can capture crucial aspects of the data in Chapters 2 and 3, albeit with assumptions that remain empirically untested and, in one case, decidedly questionable. This model cannot, however, account for the data in Experiments 4.2 to 4.5, which show that prior probabilities can induce change in the probability of the antecedent and the probability of the consequent. The next model can accommodate some of this change, while jettisoning the assumptions above.

This model is shown in Figure 4.8.

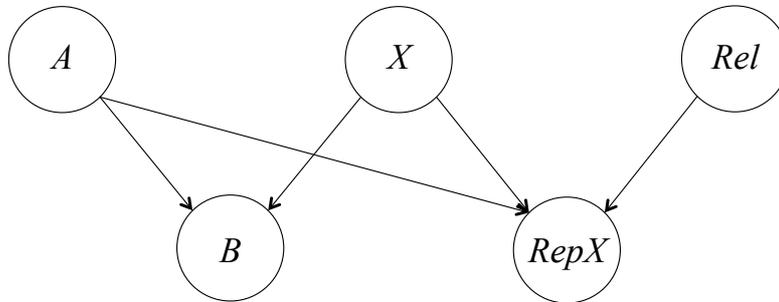


Figure 4.8. Comparison Bayesian belief network for experimental data

Structurally, the model differs from the baseline model only in the addition of an arc from *A* to *Rep X*. The probability distributions remain the same for all but *Rep X*, whose new conditional probability table is shown below.

Table 4.18. Conditional probability table for *Rep X*

<i>Conditional (X)</i>	Holds				Not			
	Reliable		Unreliable		Reliable		Unreliable	
<i>Reliability (Rel)</i>	Holds	Not	Holds	Not	Holds	No	Holds	Not
<i>Antecedent (A)</i>	Holds		Not		Holds		Not	
Report X	1	0	μ	μ	0	0	μ	μ
Not Report X	0	1	$1 - \mu$	$1 - \mu$	1	1	$1 - \mu$	$1 - \mu$

RATIONALITY, PRAGMATICS, AND SOURCES

It is assumed, here, that if the antecedent is false and the conditional holds, then a fully reliable source will not report that the conditional holds. In many cases, this assumption seems plausible. It would ordinarily be strange for the manager of a car showroom to say ‘If a car in this showroom is a Mercedes, then it’s black’ if there were obviously no Mercedes in the showroom. A slight change to the scenario will license the conditional: the showroom does not currently have any Mercedes cars, but it sometimes does. In this case, the manager could utter the conditional without oddity. But then the antecedent would not be false; the conditional would mean ‘Whenever there is a Mercedes in this showroom, then it’s black’. In other words, there are different reference classes: the current stock of cars, and all (or at least many non-specific) possible stocks of cars. We will return to this issue of the truth and falsity of antecedents below; for now, it suffices to say that, if a conditional with a false antecedent is asserted, it is considered a counterfactual, and counterfactual conditionals are typically treated separately. The assumption above, then, is plausible.

By including the extra arc, the revised model can accommodate the data in Experiments 4.2 and 4.3, which showed change to the probability of the antecedent. The nodes A and $RepX$ are now directly connected, so the prior probability of the antecedent will influence the probability of reporting the conditional, and the assertion of the conditional will alter the probability of the antecedent. The model’s behaviour can be summarized with the following equation:

$$r_0 = \frac{\mu}{(\mu + (1 - a))}$$

Here, μ is the randomization parameter for the unreliable source, and a is the prior probability of the antecedent. The probability of the antecedent increases when the prior for Rel , r , is higher than r_0 ; remains the same when r and r_0 are the same; and decreases when r is lower than r_0 . Change to the probability of the antecedent is

directly linked to the perceived reliability of the source. In contrast, both the probability of the consequent and the conditional probability always increase on assertion. Therefore, this model can, in principle, only account for the data in Experiments 4.2 and 4.3; it cannot account for the effect of the prior probability of the consequent. Furthermore, this latter effect cannot simply be captured by adding an arc between B and $RepX$, because B would not be a root node, and could not, therefore, take a prior probability³⁴. To capture this effect, we would need a more radical revision to the network, such as including a new root node with arcs to both B and $RepX$. It would be crucial, however, to independently justify the inclusion of this node to avoid the danger of over-fitting the model to the data.

The model predictions above are not testable with the current data. They require the estimation of parameters that were not manipulated or measured in the relevant experiments: namely, the randomization parameter and people's initial perception of the sources. Future work will explore how well the model fits empirical data. But both models above count as proof of principle: it is possible to capture systematic effects on at least the conditional probability and the probability of the antecedent using purely probabilistic models which flesh out the dependencies among variables. It is also striking that probabilistic models cannot straightforwardly capture all of the data from an apparently simple experiment. These difficulties

³⁴ Indeed, adding an arc from B to $RepX$ or, in addition, from A to $RepX$ simply results in an increase in probabilities ($P(A)$, $P(B)$ and $P(C)$), which is not the pattern seen in the data.

signal the complexity of learning conditionals, a complexity which seems to have been largely overlooked in the psychological literature³⁵.

It is possible, however, to judge the plausibility of the link between perceived reliability and the changes to the probability of the antecedent. Relevant evidence is to be found in the extensive corpus work of Declerck and Reed (2001). Declerck and Reed collected a small corpus of conditionals, developed hypotheses, and tested these hypotheses against data in large publicly available corpora. Their analysis was qualitative: in effect, they built a conceptual framework by systematically probing their intuitions. Their treatment of the antecedent is compatible with the intuitions and experimental data reported in this thesis. They outlined systematic variation in the truth of the antecedent, proposing the following categories (all examples are also Declerck and Reed's):

(1) *Factual Antecedent*. The antecedent is known to be true. For example, 'If I had a problem, I always went to my grandmother'. Often, the consequent refers to a past habitual action.

(2) *Non-factual Antecedent*.

a. *Neutral*. The antecedent is pure supposition. For example, 'if a woman has a history of cancer in the family, she should have herself checked out once a year'.

b. *Non-neutral*. The antecedent takes one of the following attitudes to the truth.

³⁵ This gap in the literature was repeatedly emphasized during the recent 'Learning Conditionals' workshop at Ludwig Maximilian's University, Munich, February 2nd-3rd, 2017. For discussion, see the lecture notes available at http://www.cas.uni-muenchen.de/veranstaltungen/archiv_veranstaltung/tagungen/ws_krzyzanowska_hartmann/index.html.

- i. Closed.* The antecedent is assumed to be true for the purposes of the interaction, but no commitment is necessarily made to the actual truth. For example, ‘If you didn’t do it, it must have been Fred’.
- ii. Open.* The antecedent may or may not prove true, and the conditional often refers to the future. For example, ‘If the train is late, we’ll miss our plane’
- iii. Tentative.* The antecedent presents a possibility, but one that is considered unlikely. For example, ‘If he did/were to do that, he’d be in big trouble’³⁶.
- iv. Counterfactual.* The antecedent is false. For example, ‘If I were you, I wouldn’t go there’.

Declerck and Reed (2001) use these categories to describe different types of conditionals, not to predict belief change from them. But if these categories are real to speaker-hearers, and not just corpus linguists, then they should give rise to systematic differences in belief change. This point holds most obviously for the following categories: factual-antecedent conditionals should lead to high belief in the antecedent (hence, a potential increase); and counterfactuals should lead to low belief (or, indeed, total disbelief) in the antecedent (hence, a potential decrease). Like counterfactuals, tentative-antecedent conditionals have a morpho-syntactic cue (the verb form), in this case suggesting that the antecedent, though possible, is unlikely. These conditionals should also lead to a low belief in the antecedent. The remaining categories are harder to interpret. Both closed and open conditionals do

³⁶ This example seems equivalent to ‘If he DOES do that, he’ll be in big trouble.’

not, on Declerck and Reed's account, require a commitment to the truth or falsity of the antecedent. It is plausible, though, that, here too, the probabilities may shift, depending on context.

The framework above suggests a research program to explore belief change from different types of conditional, as these different types may prompt different patterns of belief revision. Notably, though, Declerck and Reed (2001) do not refer to source reliability; they refer, rather, to context and linguistic cues such as the verb form. Indeed, Declerck and Reed give many examples which have markedly different interpretations even with little or no context for the judgment of source reliability. This point raises doubts about the proposed mechanism by which prior probabilities are said to influence belief change. The diversity of examples also prompts the suggestion that a single model may not fit all types of indicative conditional. Instead, a series of models might be required.

Before we leave the topic of modelling testimonial conditionals, it will be useful to summarize what the foregoing discussion reveals about learning a conditional. The experimental data suggest that, when we learn a conditional, we increase our judgment of the conditional probability. They also suggest that, under certain circumstances, we modify our judgments of the probability of the antecedent and the probability of the consequent. Considering the models reveals other possibilities. In both models, when *RepX* is set to true, then probabilities change simultaneously, not just at *X*, *B* (unless, in the first model, certain conditions hold), and (in the second model) *A*, but also at *Rel*. In other words, the models suggest that, when we hear a conditional, we systematically revise our beliefs about the reliability of the source. These model predictions should be tested in future work.

4.7.2 Probabilities and Beyond

This thesis has focused on a small set of probabilities: the probability of the antecedent, the probability of the consequent, and the conditional probability. Although these probabilities seem closely implicated in testimonial conditionals, recent research suggests that people expect conditionals to come with more information. Intuitively, conditionals are strange if there is no link between antecedent and consequent, quite apart from the conditional probability. For example:

If the capital of France is Paris, then the Earth is spherical.

Since we are certain about the consequent, then $P(\text{Spherical}|\text{Paris}) = 1$, but this conditional is odd and very likely unassertable. One way to explain the oddity is to argue that the antecedent and consequent must be relevant. Relevance is understood, here, in terms of the relationship between two conditional probabilities, $P(\text{Consequent}|\text{Antecedent})$ and $P(\text{Consequent}|\text{NotAntecedent})$ (Skovgaard-Olsen, 2016; Skovgaard-Olsen et al., 2016b; Skovgaard-Olsen, Singmann, & Klauer, 2016a):

Positive Relevance:

$$P(\text{Consequent}|\text{Antecedent}) > P(\text{Consequent}|\text{NotAntecedent})$$

Irrelevance:

$$P(\text{Consequent}|\text{Antecedent}) = P(\text{Consequent}|\text{NotAntecedent})$$

Negative Relevance:

$$P(\text{Consequent}|\text{Antecedent}) < P(\text{Consequent}|\text{NotAntecedent})$$

On this view, when we hear a conditional, we first ask whether the antecedent is a sufficient reason for the consequent – that is, we ask whether there is positive relevance - then we evaluate the conditional probability. If antecedent and

consequent are positively relevant, then the probability of the conditional is simply the conditional probability. But if antecedent and consequent are irrelevant or negatively relevant, a penalty applies: the probability of the conditional is, consequently, lower than the conditional probability. Experimental data support this account (Skovgaard-Olsen, 2016; Skovgaard-Olsen et al., 2016a, 2016b).

This relevance-based approach implies that, when we learn a conditional, we learn that antecedent and consequent are positively related, that is, that $P(\text{Consequent}|\text{Antecedent}) > P(\text{Consequent}|\text{NotAntecedent})$. Collins et al. (2017) report another data set, as part of the study reported here, which included measurements of both conditional probabilities³⁷. These data show that both assertion and expertise predict relevance: assertion shows a significant increase in relevance from null to single and null to multiple conditions (though a non-significant decrease from single to multiple); and expertise shows a significant increase from null to inexpert, null to expert, and inexpert to expert conditions.

It is not entirely clear how to characterize this relevance requirement. For instance, relevance might be fundamental to the conditional (Skovgaard-Olsen, 2016; Skovgaard-Olsen et al., 2016a, 2016b); it might, alternatively, arise merely out of discourse coherence, or apply to connectives more generally (Cruz, Over, Oaksford, & Baratgin, 2016). Suggestive data come from two studies by Krzyżanowska, Collins and Hahn (2017a,b). These studies explored whether probabilistic relevance is distinguishable from mere discourse coherence and whether conditionals and conjunctions have the same assertability conditions. The

³⁷ This data set is not reported here, for brevity's sake. The study used the methodology of the original and replication study, and provided a further replication of those data, while collecting additional measurements.

resultant data suggested that, while simple conversations required only the same topic to be coherent, the same information expressed in conditionals required probabilistic relevance on top of topic coherence. Moreover, conditionals and conjunctions diverged in their assertability conditions. These data suggest, then, that probabilistic relevance is a fundamental and distinctive component of conditionals.

4.7.3 The meaning of the conditional

Since we have discussed the full set of data and attempts to model them using Bayesian belief networks, we are now in a better position to judge the implications for the meaning of the conditional. As we have seen, the data directly contradict predictions of two accounts: the material conditional; and a simple Bayesian network coupled with the Kullback-Leibler divergence. These accounts imply that, when a conditional is asserted, the probability of the antecedent should decrease. Chapter 2 found no evidence for this effect, and this chapter has shown that the probability of the antecedent reliably increases when the prior probability is low. The data are consistent with the suppositional theory, in the weak sense that they do not contradict the theory but are not predicted by it. If the suppositional theory is supported by modelling with Bayesian belief networks, then it achieves better, but by no means perfect, predictive success. The data are also consistent with the notion that conditionals require some link between antecedent and consequent, be it relevance or some inferential relationship. But this is, again, a relatively weak consistency: relevance and inferentialism do not predict systematic changes to the

probability of the antecedent or the probability of the consequent. The data do not seem, then, to fit neatly with any particular theory of the conditional.³⁸

It is tempting, here, to invoke pragmatics: the experiments are, after all, pragmatic contexts. This chapter has already speculated on a pragmatic account. At work, in this account, are two guiding pragmatic maxims: the Maxim of Relevance, and the Maxim of Quantity. Take, first, the case when the prior probability of the antecedent is low. Here, the Maxim of Relevance operates: the speaker would not be uttering a conditional unless there was a substantial chance that the antecedent would obtain. When, in contrast, the prior probability of the antecedent is high, the Maxim of Quantity operates: the speaker is uttering a weaker statement than justified by the hearer's prior; the speaker may then be implicating that there is insufficient evidence for the stronger statement. This account can be extended to the consequent: a conditional acquires relevance by having instances, hence some possibility of the consequent being true (Sperber, Cara, and Girotto, 1995). But if the consequent is known to be true, or highly probable, then a conditional may be weaker than the justifiable information content.

Although the suggestions above are plausible, they are *ad hoc*. They are arguably an example of the pragmatic waste bin. Why so? The experiments do not – and, indeed, were not intended to – tease apart semantics and pragmatics. It is one thing to say that the leading semantic theories do not account for the data; it is quite another to say that the effects are, therefore, pragmatic. The experiments were intended to explore belief change from testimonial conditionals and, from the point

³⁸ The listed theories are not the only accounts of the conditional. However, they are particularly influential at present. A principal alternative, the Stalnaker conditional, does not seem to make clear predictions without extra assumptions (Douven, 2012).

of view of testimony, what is crucial is the information communicated and accepted, not whether that information is communicated semantically or pragmatically.

Nevertheless, it is an interesting and theoretically important question for linguistics, the philosophy of language, and the psychology of reasoning, whether the effects are semantic or pragmatic.

Since, to repeat, the present experiments were not designed to tease apart semantics and pragmatics, they used a dependent measure that is too crude for that purpose. A variety of dependent measures feature in the experimental literature on the conditional. Within the psychology of reasoning, the measures are overwhelmingly ratings scales: ratings of truth, probability, believability, assertability, or acceptability. The first three ratings scales – truth, probability, and believability – are often treated as accessing semantics (on truth and believability, see Douven & Krzyżanowska, 2017; on probability, see Evans & Over, 2004, or Over, 2016); the remaining ratings scales, as accessing pragmatics. However, it is far from clear that any scale isolates semantics or pragmatics. Take, for example, the following sentence from Douven and Krzyżanowska (2017):

Princess Diana died in a car accident and she divorced Prince Charles.

A textbook account is that such sentences are semantically true but pragmatically false because of an order implicature: the sentence implicates, nonsensically, that the death happened before the divorce. A hearer could reasonably respond ‘Yes, but not in that order’, or ‘No, she divorced Prince Charles *before* she died in a car accident.’ There is, moreover, no *a priori* reason to discount the possibility that hearers would judge the probability of the sentence above to be different from a sentence with a more natural ordering. Plausibly, the sentence’s believability will also depend on the

ordering (Douven, 2010). And semantic factors seem likely to contribute to the assertability and acceptability of the statement³⁹.

Although ratings scales do not pick out semantics or pragmatics uniquely, we could nevertheless argue that some scales are more sensitive to one or the other, use all scales together, and then interpret differences among these scales. However, recent empirical data suggest that there are, in fact, no clear differences in data among ratings scales. Douven and Krzyżanowska (2017) presented participants with a range of sentences which were semantically true but pragmatically false, such as the Princess Diana example above. They asked participants to rate the truth, assertability, or believability of the sentences. The resultant data showed no difference among the ratings scales, either in a two-option forced-choice task or with Likert-style scales. These data cast further doubt on the use of ratings scales to separate semantics and pragmatics.

A more promising strategy is to borrow methods from psycholinguistics. In psycholinguistics it is common to use both offline and online measures: offline measures include grammaticality or truth-value judgments; online measures include reaction times, reading times in self-paced reading tasks, eye movements, and event-related potentials (Harley, 2013). Online measures can provide crucial information about factors such as the time course of processing: when certain types of information become available. There is, for instance, a vigorous debate about when the word ‘some’ is enriched to mean ‘not all’ (e.g. Breheny, Ferguson, & Katsos, 2013; Breheny, Katsos, & Williams, 2006; Grodner, Klein, Carbary, & Tanenhaus,

³⁹ Krzyżanowska, Collins, and Hahn (2017a, b) find differences in assertability when conversational contexts are kept constant. These differences suggest that ratings of assertability are sensitive to semantics.

2010). Information about time course can also bear on whether a phenomenon is semantic or pragmatic. Semantic information tends to be instantly available, whereas pragmatic information takes additional processing, and may therefore be somewhat delayed⁴⁰ (A. Stewart, personal communication, October 7, 2016). Such data are not, of course, watertight (see, e.g., Douven & Krzyżanowska, 2017), but they are, nevertheless, suggestive.

We bring this chapter to a close by considering how the data reported in Chapters 2 to 4 bear on the triangular scheme. These chapters have focused on the link between pragmatics and sources. The third component, rationality, has remained largely in the background, although it is implicit in any discussion of testimony. We can draw out the role of rationality by reconsidering Crupi's model for detecting (ir-)rationality in cognitive experiments. In this model, a crucial component is the match between experimenters' and participants' understanding of the premises. According to Crupi's model, we cannot pronounce on rationality until we have a theory of how participants understand the premises. But Chapters 2 to 4 have argued, conceptually and empirically, that cognitive science lacks a good theory of testimonial conditionals. Since testimonial conditionals feature in any experiment on conditional reasoning, we must remain cautious about pronouncing on the (ir-)rationality of people's conditional reasoning.

This abstract point can be made clearer with an example. Take the debate on centring inferences: inferences of the type 'P and Q. Therefore, if P then Q.' and 'P. Q. Therefore, if P then Q'. Although such inferences seem counter-intuitive, they are

⁴⁰ There are competing models of utterance processing, some of which do not take literal meaning to be processed faster than non-literal meaning in all cases. There is

licensed by major theories of the conditional, including the material interpretation (Grice, 1989; Jackson, 1979), the Stalnaker conditional (Stalnaker, 1975), Mental Models Theory (Johnson-Laird & Byrne, 2002), and the suppositional theory (Adams, 1975; Edgington, 1995; Evans & Over, 2004). Centring is not licensed by theories which assume that conditionals require relevance or explanatory relations between antecedent and consequent (Douven, 2015; Krzyżanowska, 2015; Skovgaard-Olsen, 2016; Skovgaard-Olsen et al., 2016b). Empirical studies produce contradictory results, suggesting that people endorse (Cruz, Baratgin, Oaksford, & Over, 2015) and do not endorse centring (Skovgaard-Olsen et al., 2016a).

Is centring a rational inference? To ask this question amounts to asking how people interpret conditionals in an inference task: in a given task, do people interpret a conditional as requiring a relevance or explanatory relation, or will the conditional probability suffice? To predict people's inferences in context and judge their rationality, we will need to predict people's interpretations in context. Chapters 2 to 4 suggest that we are some way off having a robust theory of the semantics and pragmatics of the conditional. Developing such a theory will be an essential complement to the rich existing literature on conditional reasoning.

reasonable evidence, though, that properly inferred meanings take longer. For discussion, see, for example, Giora (1999, 2002, 2009)

5 Framing and utility conditionals

Chapter 1 introduced a triangular scheme, arguing that rationality, pragmatics, and sources are intimately related and that we benefit from studying them together. Chapters 2 and 4 focused on the link between pragmatics and sources in testimonial contexts, demonstrating how assertion and expertise influence people's judgments of conditionals' information content. This chapter focuses on the link between pragmatics and rationality, using goal framing as a case study to illustrate what pragmatics can tell us about rationality and rationality about pragmatics. This chapter will, first, re-introduce framing as a general topic, before introducing goal framing as the research focus. It will argue that utilities draw together pragmatics and rationality: people use utilities to interpret conditionals, and to assess conditional arguments; how they do this directly impacts on theories of framing, since key framing effects rely on the conditional.

As we saw in Chapters 1 and 2, framing is a key phenomenon in the psychology of rationality. Framing effects are taken as important evidence that human judgment and decision making is suboptimal. These effects take different forms, which vary in complexity. In each case, though, people make different judgments or decisions on the basis of different but supposedly equivalent descriptions, hence (again, supposedly) violating the principle of description invariance (Tversky & Kahneman, 1981). Much rides on the notion of equivalence: although non-equivalent frames are also researched, they do not have such obvious implications for rationality (Druckman & McDermott, 2008).

5.1.1 Prospect Theory and the Asian Disease Paradigm

The most influential account of framing is due to Tversky and Kahneman (1981). The account runs thus. Framing occurs when people are faced with a decision problem. Decision problems are defined by the available options, the attendant outcomes, and the conditional probabilities of these outcomes given the available options. Decision makers approach decision problems by, first, establishing a decision frame (a stage often called ‘editing’) and then evaluating the options. Decision frames are ‘[the] decision maker’s conception of the acts, outcomes, and contingencies associated with a particular choice’ (Tversky & Kahneman, 1981, p.453). Which frame is chosen is a highly contextual matter. The frame is ‘controlled partly by the formulation of the problem and partly by the norms, habits, and personal characteristics of the decision-maker (1981, p.453).’ A key aspect is whether outcomes are formulated as gains and losses. This formulation affects how people evaluate options. People evaluate options against a neutral but movable reference point. Relative to this reference point, people are risk-averse with gains and risk-seeking with losses (for exposition in terms of the valuation function see, e.g., Kahneman & Tversky, 1979; Tversky & Kahneman, 1981).

The above account applies most straightforwardly to risky-choice framing. A clear example is the Asian disease paradigm discussed in Chapter 1. It will be useful to briefly repeat the description of the paradigm and the prospect-theoretic explanation here. In the Asian disease paradigm, participants see the following (or equivalent) materials:

Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific

estimate[s] of the consequences of the programs are as follows:

[Usually seen by one group]

If Program A is adopted, 200 people will be saved.

If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.

[Usually seen by another group]

If Program C is adopted, 400 people will die.

If Program D is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.

(Tversky & Kahneman, 1981, p. 453)

Programs A and B appear under gain framing: the frame is ‘lives saved’, which sets the reference point at 600 deaths. Any life saved is a gain relative to this reference point. Since gains prompt people to be risk-averse, people prefer the certain option, Program A. Programs C and D, in contrast, appear under loss framing: the frame is ‘will die’, which sets the reference point at 0 deaths. Any death is a loss relative to this reference point. Since losses prompt people to be risk-seeking, people prefer the risky option, Program D. With these (and equivalent) materials, this preference reversal replicates well (Kühberger & Tanner, 2010), though it is unclear whether the preference reversal is robust to variations in the specified information (Chick et al., 2016; Kühberger & Tanner, 2010; Mandel, 2014).

5.1.2 Towards goal framing

This chapter focuses on a different but related framing effect, known as goal framing. Although there are subtle differences between the effects, goal framing, too,

RATIONALITY, PRAGMATICS, AND SOURCES

is taken to challenge rationality (Levin, Gaeth, Schreiber, & Lauriola, 2002; Levin et al., 1998). Goal frames are frames such as the following:

Positive frame: If you decide to get HIV tested, you may feel the peace of mind that comes with knowing about your health.

Negative frame: If you do not decide to get HIV tested, you may not feel the peace of mind that comes with knowing about your health.

(Apanovitch et al., 2003)

Goal framing differs from ‘Asian disease’ framing in how it applies to actions. In both cases, there are two choices. In the ‘Asian disease’ paradigm, the choice is between two alternative actions; framing occurs if there are preference reversals. In goal framing, the choice is to perform or not perform a particular action (Levin, Schneider & Gaeth, 1998). Goal frames, that is, argue in favour of an action either by stressing the positive consequences of undertaking an action or by stressing the negative consequences of not undertaking it (Levin, Schneider & Gaeth, 1998). In goal framing, framing occurs if one frame or the other prompts higher ratings of persuasiveness, attitude, or intention to perform, or prompts more people to actually undertake the action (Gallagher & Updegraff, 2012).

The types of framing also differ with respect to both probabilities and utilities. These differences can be understood more precisely by considering the structure of the conditionals themselves. Take, first, the component utilities. The typical ‘Asian disease’ conditional has the form ‘If Program A is adopted, then 200 people will be saved.’ The antecedent - ‘Program A is adopted’ – has no obvious utility; there is no information about Program A itself, other than the outcome specified in the consequent, ‘200 people will be saved’. Only the consequent has clear utility. In contrast, the typical goal-framing conditional has the form ‘If you decide to get HIV

RATIONALITY, PRAGMATICS, AND SOURCES

tested, you may feel the peace of mind that comes with knowing about your health'. Here, the antecedent does have (dis)utility: although the precise value will vary subjectively, HIV tests are associated with at least mild inconvenience and discomfort. The consequent – 'you may feel the peace of mind...' – again specifies an outcome with utility. There are also differences in the conditional probabilities. In 'Asian disease' conditionals, the conditional probabilities are specified. This specification is implicit for Programs A and C: presumably, for example, $P(200 \text{ saved} | \text{Program A}) = 1$. It is explicit for Programs B and D: for instance, $P(600 \text{ saved} | \text{Program B}) = 1/3$. In goal-framing conditionals, there are typically no probabilities (Rothman & Salovey, 1997).

Despite these differences, goal-framing research has come to test the same prospect-theoretic predictions as 'Asian disease' framing (Rothman & Salovey, 1997). This process started after the failure of an early theory: that, because people are loss averse and have a negativity bias, loss frames should be more effective across the board (Meyerowitz & Chaiken, 1987). An early goal-framing study explored this prediction in communication about breast self-examination (Meyerowitz & Chaiken, 1987). Participants read pamphlets promoting breast-self examination which used gain or loss framing or no arguments. For example,

By (not) doing BSE now, you can (will not) learn what your normal, health breasts feel like so that you will be better (ill) prepared to notice any small, abnormal changes that might occur as you get older.

(Meyerowitz & Chaiken, 1987, p. 504)

Participants who read loss-framed pamphlets showed more positive attitudes, intentions and behaviours with respect to breast-self examination. Subsequent research has produced less clear results. Some studies show no effect (Lalor &

RATIONALITY, PRAGMATICS, AND SOURCES

Hailey, 1989); others suggest moderation by variables such as efficacy, personal involvement, threat, or mood (Block & Keller, 1995; Rothman, Salovey, Antone, Keough, & Martin, 1993; Wegener, Petty, & Klein, 1994). Meta-analyses demonstrate, however, that, overall, loss frames are not more persuasive than gain frames (O’Keefe, 2011; O’Keefe & Jensen, 2006, 2007, 2009).

Apparent inconsistencies among studies led Rothman and Salovey (1997) to reanalyze goal framing and introduce new theoretical distinctions to attempt to explain the divergent results⁴¹. They introduced two principal innovations: a distinction between advocated behaviours; and a distinction between phrasings. Take, first, the distinction in behaviours. Different behaviours, they argued, differ systematically in risk, understood in a specific and unusual way:

‘the risk associated with a behavioural alternative usually cannot be defined in terms of the actual likelihood of a particular outcome. Instead, *risky* reflects the subjective perception that to perform a behaviour may involve an unpleasant outcome’

(Rothman & Salovey, 1997, p. 5)

‘Risk’, here, would perhaps be better understood as utility. Nevertheless, Rothman and Salovey used this notion of risk to justify a (somewhat questionable) move to Prospect Theory, according to which gain frames prompt risk aversion and loss frames prompt risk-seeking behaviour. Prevention behaviours include eating healthily, exercising, and taking preventative medicines. Prevention is taken to be

⁴¹ O’Keefe (2011) has suggested that the loss aversion explanation– and the subsequent prevention/detection distinction – survived so long because researchers have favoured narrative reviews, where conclusions are based on the presence or absence of a statistically significant result. It was, on his view, the switch to meta-analytic methods which cast doubt on both accounts.

RATIONALITY, PRAGMATICS, AND SOURCES

low in risk: that is, the ‘relatively certain’ outcome is positive (Rothman & Salovey, 1997, p. 10). Detection behaviours include health screening, and are taken to be high

in risk: there is a chance of a negative outcome (detecting an illness), at least in the short term. It is the short-term consequences, on this account, that are crucial. The second principal distinction is in phrasing: the general terms ‘gain’ and ‘loss’ can be realized in the follow ways:

Gain: Attain Gain If you get HIV tested, you may feel the peace of mind which comes with knowing about your health.

 Avoid Loss If you get HIV tested, you may not suffer the anxiety of not knowing about your health.

Loss Avoid Gain If you don’t get HIV tested, you may not feel the peace of mind which comes with knowing about your health.

 Attain Loss If you don’t get HIV tested, you may suffer the anxiety of not knowing about your health.

These different phrasings could, in principle, affect persuasive outcomes: the ‘attain gain’ and ‘attain loss’ frames may have simpler representations than the ‘avoid loss’ and ‘avoid gain’ frames (Brendl, Tory, & Lemm, 1995).

5.1.3 Goal framing: the empirical data

How do these distinctions map onto the data? Although Rothman and Salovey’s (1997) distinctions are plausible, they lack clear evidence. One meta-analysis showed a small but significant advantage of gain frames for prevention behaviours, but this was largely driven by a larger effect in data on dental health (O’Keefe & Jensen, 2007). Another meta-analysis showed a small but significant advantage of loss frames for detection behaviour, but this was largely driven by a

larger effect in data on breast-cancer detection (O’Keefe & Jensen, 2009). More problematically, a subsequent large-scale meta-analysis has shown a small but significant advantage of gain frames for prevention behaviours, especially for skin-cancer prevention, smoking cessation and physical activity, but no significant difference in data on attitudes, intentions or detection behaviours (Gallagher & Updegraff, 2012). The data on phrasing is similarly inconclusive. Few studies have used the necessary manipulations, and those studies produce contradictory results: Detweiler, Bedell, Salovey, Pronin, and Rothman (1999) found no difference between the types of gain and loss frames; Nan (2007), however, only found framing when using ‘avoid loss’ and ‘attain loss’ frames (see, also, O’Keefe & Jensen, 2006, 2007, 2009).

Why do the data not support framing? We will consider three options here: that Prospect Theory simply does not apply to the standard approach; that any framing account is based on a false assumption of equivalence between frames; and that the literature has not systematically manipulated crucial predictors.

Take, first, the possibility that Prospect Theory simply does not apply to the standard goal-framing task. The plausibility of this suggestion hinges on two points: the interpretation of ‘risk’, and the options available. The definition of ‘risk’ in goal-framing research seems to have evolved. As we have already seen, Rothman and Salovey (1997) interpreted ‘risk’, in effect, as utility: this interpretation is at play in the distinction between prevention and detection behaviours. It is by no means clear that Prospect Theory applies with this interpretation of risk. However, by Apanovitch et al. (2003), the interpretation shifted to focus on perceptions of certainty, an interpretation which is more congenial to Prospect Theory. Apanovitch et al. claimed, without citing empirical data, that prevention behaviours are typically

RATIONALITY, PRAGMATICS, AND SOURCES

viewed as having more certain outcomes than detection behaviours. Although their evidence for this claim is unclear, their experiments did, indeed, suggest that perceived certainty is an important factor: perceived certainty led to an advantage for gain frames in HIV testing. Their focus on certainty does not, however, seem to have been widely shared (for exceptions, see Block & Keller, 1995; Gerend & Cullen, 2008; Meyers-Levy & Maheswaran, 2004; O'Connor, Ferguson, & O'Connor, 2005; Schneider et al., 2001).

The emphasis on perceived certainty brings goal framing closer to Prospect Theory, but not close enough to derive reasonable predictions. The culprit, here, is the number of options available to participants. Recall that in the Asian Disease Paradigm, where Prospect Theory straightforwardly applies, participants choose between two options: a certain and a risky option. It is this choice that framing acts on: gain frames prompt the choice for the certain option; loss frames, for the risky option. In contrast, in gain framing participants choose between compliance and non-compliance with the advocated behaviour, but these options need not differ in certainty (O'Keefe & Jensen, 2007). To borrow O'Keefe and Jensen's examples:

- 1.a. If I brush my teeth regularly, I'll almost certainly avoid cavities.
- 1.b. If I don't brush my teeth regularly, I might avoid cavities or I might not.
- 2.a. If I get the flu jab, I might or might not get the flu.
- 2.b. If I don't get the flu jab, I might or might not get the flu.

1.a. and 1.b. differ in certainty; 2.a. and 2.b. do not. Note that the behaviours mentioned are prevention behaviours. These variations suggest that there is limited scope for general predictions from Prospect Theory; that they are variations within a single class of behaviour, prevention behaviours, is further evidence against a helpful role for the prevention/detection distinction.

Let us temporarily assume, for the sake of argument, that the notions of risk and the behavioural options are no obstacle to applying Prospect Theory to goal framing. We, nevertheless, encounter another reason for the failure of prospect-theoretic predictions: these predictions are based on a false assumption that the frames are equivalent. Earlier in this chapter, we saw that equivalence is crucial to the interpretation of framing. The goal-framing literature assumes that positive and negative frames convey equivalent information (for discussion, see Corner & Hahn, 2010). That is, the literature assumes that ‘If P then Q’ and ‘If not P then not Q’ are equivalent. The question naturally arises ‘Which standard should we use for equivalence?’ A natural candidate is classical logic: according to classical logic, if the frames are equivalent, then they should entail each other. But classical logic prohibits the inference ‘If P then Q; therefore if not P then not Q’ and *vice versa*. Classical logic does licence the inferences ‘If P then Q; therefore if not Q then not P’ and (simplifying somewhat) ‘If not P then not Q; therefore if Q then P’. But this inference is not useful in the appropriate contexts: there is no persuasive value in inferring ‘If your risk of cancer does not decrease, then you are not wearing sunscreen’ from ‘If you wear sunscreen, then your risk of cancer decreases.’ An alternative standard is probability. But with probabilities, too, frames are not necessarily equivalent. While the conditional probabilities $P(\text{Consequent}|\text{Antecedent})$ and $P(\text{Not Consequent} | \text{Not Antecedent})$ may contingently be the same, they do not have to be. That is, they do not formally constrain each other. Considering these inferences alone is enough to reject equivalence. The content of the conditionals may increase the non-equivalence still further. For instance, there is no reason to assume that hearers will assess the relevant utilities equivalently under positive and negative frames, from the point of

view either of estimating the speaker's assessment of the utilities or of making their own subjective estimates.

The non-equivalence of frames may account for the failure to support predictions from Prospect Theory. It may also explain the lack of systematic differences between positive and negative frames: the literature has simply not manipulated the appropriate predictors. One key predictor is the conditional probability. This thesis has reviewed an extensive literature which suggests that the conditional probability is central to the information content of the conditional. It has also presented empirical data in support of this suggestion. We should strongly expect, then, that the conditional probability is important to these conditional frames. As we have seen, though, probabilities have largely been set aside by goal-framing researchers, because 'the risk associated with a behavioural alternative usually cannot be defined in terms of the actual likelihood of a particular outcome' (Rothman & Salovey, 1997, p. 5). The decision is based on questionable reasoning. Whether or not such risks can be defined in terms of the objective likelihood (and, for many health behaviours, such information may well be available), they can be estimated subjectively, and participants' subjective estimates could be manipulated using experimental materials. This neglect of probabilities, then, is unnecessary as well as problematic.

Important though probabilities are, this chapter will focus on a manipulation which is closer to the existing literature: namely, utilities. Doing so will allow us to consider the relationship between pragmatics and rationality from a fresh perspective for this thesis. This chapter will focus on the two utilities which are directly implicated in goal frames: the utility of the antecedent and the utility of the consequent.

RATIONALITY, PRAGMATICS, AND SOURCES

Neither utility seems to have been explored satisfactorily in empirical research. Take, first, the utility of the antecedent. Some researchers have assumed that the antecedent is inert with respect to utilities: that the antecedent is a good thing to do, so its utility can be ignored in experiments (Levin et al., 1998). This assumption is intuitively unsatisfactory, since antecedent behaviours can clearly vary in their intrinsic utility. Compare, for instance, two prevention behaviours: taking regular medicines to prevent heart disease, and undergoing surgery to prevent breast cancer. While both the antecedent behaviours may lead to positive consequences, they differ starkly in their intrinsic utility; this difference seems intuitively likely to contribute to the persuasiveness of an argument. Other researchers have explored the utility of the antecedent through the prevention/detection distinction: this relies on another assumption, in early accounts of this distinction, that prevention behaviours have positive utility (or are, in their terms, low risk) and detection behaviours have negative utility (or are, in their terms, high risk) (Rothman & Salovey, 1997; Rothman et al., 1993). But this manipulation is not systematic: prevention behaviours can differ widely from each other in utility, as can detection behaviours. Still other researchers collect participants' utility judgments indirectly, through measures of attitude towards the health behaviour in question (e.g. Uskul, Sherman, & Fitzgibbon, 2009; van Assema, Martens, Ruiters, & Brug, 2001). But these data are natural variations among participants, not systematic manipulations. It may, of course, be unfeasible to manipulate antecedent utilities in a task whose materials simulate a health campaign: the utilities are constrained by the facts of the matter. But, as we will see below, more abstract experimental tests are both possible and desirable: they allow us to investigate the mechanisms underlying the persuasiveness of these frames, mechanisms which can later be explored in more naturalistic tasks.

RATIONALITY, PRAGMATICS, AND SOURCES

Matters are somewhat different for the utility of the consequent. There is a fair amount of relevant evidence which has been generated to test other hypotheses. O’Keefe (2013) argued that a wide range of goal-framing studies (indeed, persuasion studies more generally) can be re-interpreted as exploring the desirability – hence, utility – of the consequences. Many studies have, for instance, sought moderating variables for goal framing, typically individual-difference variables such as people’s concern for the future or their regulatory focus (for discussion, see O’Keefe, 2013). For example, Orbell and Hagger (2006) investigated how people’s concern for future consequences affected their responses to messages about diabetes screening. People who showed low concern were more persuaded by messages focusing on short-term positive, and long-term negative, consequences. People who showed high concern were, conversely, more persuaded by messages focusing on short-term negative, and long-term positive, consequences. Cesario, Grant, and Tory (2004) found similar persuasive differences according to regulatory focus. They presented arguments for an after-school program which described the program as either promoting success or preventing failure. Persuasiveness was predicted by the fit with participants’ regulatory focus: participants who were promotion-focused (motivated more by achieving gains) were more persuaded by arguments emphasizing achievement of gains (promoting success); participants who were prevention-focused (motivated more by avoiding losses) were more persuaded by arguments emphasizing avoidances of losses (avoiding failure). The data from these studies can be interpreted as suggesting that frames with desirable consequents are more persuasive. Indeed, O’Keefe suggests that these data add further to support to what he calls the single best supported empirical generalization about persuasion: arguments with desirable consequences are likely to be persuasive. O’Keefe’s analysis suggests, then, that the

utility of the consequent, like that of the antecedent, should be systematically manipulated to explore the persuasiveness of goal framing. Such systematic manipulation does not feature in the experimental literature.

Considering utilities in this way argues for taking a somewhat simpler approach to goal framing than has been taken so far in the literature. To understand the frames, we can abstract away from complex simulations of health campaigns and, instead, systematically manipulate components of the frames themselves. The research question will be whether the effectiveness of a frame depends on the subjective utility of the antecedent and the subjective utility of the consequent. As we will see, considering the utilities of the frames will reveal subtle phenomena that were otherwise hidden in more complex designs.

5.1.4 Towards an experimental paradigm: utilities in argumentation and pragmatics

While utilities do not feature systematically in the goal-framing literature, they have received far more attention in two parallel literatures: in the study of arguments from consequences, and in pragmatics. These literatures treat inextricably linked, though not identical, phenomena. They also suggest an experimental paradigm.

As we have already seen, a large body of research on persuasion suggests that people are sensitive to consequences and, in particular, supports a role for the utility of the consequent. A parallel literature, in pragmatics, has manipulated the utility of the antecedent and the utility of consequent. This literature focuses on inductive speech acts: acts of inducing or persuading someone to do something (Bonnefon, 2012). It may seem surprising to think of inducement or persuasion as a pragmatic phenomenon. However, speech acts have long been viewed from the

RATIONALITY, PRAGMATICS, AND SOURCES

perspective of their effects – or, in Austin's (1965) term, perlocutions - including their persuasive effects. Indeed, the very aim of speech acts is to have an effect on someone, be it on their feelings, thoughts, or actions (Austin, 1965; Marcu, 2000). Inducive speech acts are relevant and assertable in virtue of the utilities of antecedent and consequent and the conditional probability of the consequent given the antecedent (Evans, Neilens, Handley, & Over, 2008; Ohm & Thompson, 2004). For instance, tips and promises are effective when the cost of the antecedent is low, the benefit of the consequent is high, and the conditional probability is high; warnings and threats are effective when the benefit of the antecedent is low, the cost of the consequent is high, and the conditional probability is high (Evans et al., 2008).

A broader account has been laid out by Bonnefon in his theory of utility conditionals (Bonnefon, 2009, 2012, 2016, Bonnefon et al., 2011, 2013; Bonnefon & Sloman, 2013). This account has both rational and pragmatic components: it is intended to integrate both work oriented more towards rationality (Corner, Hahn, & Oaksford, 2011) and work oriented more towards pragmatics (Evans et al., 2008). At the heart of the theory is a formal device, the utility grid, presented in annotated form below:

If	x	u	y
Then	x'	u'	y'
	Actor	Utility	Target

To illustrate, consider the conditional 'If I let you get away with it, then my boss will fire me':

If	s	+	h
Then	e	-	s
	Actor	Utility	Target

RATIONALITY, PRAGMATICS, AND SOURCES

The actor in the antecedent is the speaker, *s*; the antecedent is about the speaker performing an action with positive utility (+) for the hearer, *h*. The actor in the consequent is a third party, the boss, who is neither the speaker nor the hearer, and is represented as *e*. The consequent is about the boss performing an action with negative utility (-) for the speaker. Any entity can be an actor or target; for events without an agent, then the neutral agent *w*, for the World, can be assigned. Utilities are represented using a simple ordinal scale: --, -, 0, +, ++, on the justification that a continuous scale is psychologically implausible (Bonneton, 2009). It is probably more usual, in fact, to assume a continuous scale (Read, 2007), but there will be no need, in this thesis, to take a stand on this issue.

Alongside the utility grid is a set of folk axioms: assumptions that ordinary reasoners are presumed to make about human decision making. Reasoners are taken to assume that people generally take actions that increase their own personal utility and avoid actions that decrease it; that other people should take actions that increase their (the reasoner's) utility, and avoid actions that decrease it; and that people take actions that increase other people's utility as long as doing so does not decrease their own personal utility, and avoid actions that decrease other people's utility as long as doing so does not increase their own personal utility (Bonneton, 2009).

Armed with the utility grid and the folk axioms, Bonneton can account for a wide range of data on conditionals with valued antecedents and consequents. His framework can be used to predict participants' inferences about a speaker's intention or about the likely behaviour of one of the mentioned people (actors or targets) (Bonneton, 2012). These successful predictions are remarkable given that the theory includes a small set of ordinal utilities: the more so because, for technical reasons,

RATIONALITY, PRAGMATICS, AND SOURCES

the theory eschews probabilities⁴² (Bonnefon, 2009). Bonnefon's framework can also be used to predict the persuasiveness of utility conditionals. For present purposes, the most relevant example is goal frames:

'If you use sunscreen, your risk of skin cancer will decrease'.

(h – h; w ++ h)

It is plausible to assume, here, that, for many people, applying sunscreen is mildly inconvenient, hence has mild disutility, and that one's risk of skin cancer decreasing has considerable utility. For such people, the aggregate utility (U(Antecedent) + U(Consequent)) is mildly positive (+). Assuming that people favour actions which increase their personal utility, then this conditional should be persuasive. For any individual, though, the actual persuasiveness will depend on their subjective utilities. In this simple case, probabilities may not be relevant: the relationship between antecedent and consequent seems deterministic, although there will presumably be a distribution over the sizes of risk decreases. This simple analysis is trivial from the point of view of the theory of utility conditionals, but nevertheless offers an advance over the goal-framing literature, because both the key utilities are considered.

As we have seen, goal framing requires a pair of conditionals, generally of the form 'If (not) P then (not) Q'. The question naturally arises of how to represent the negative frame. This apparently straightforward question does not meet with a simple answer in the literature: neither the argumentation literature nor the pragmatics literature appears to have addressed this issue. Jean-François Bonnefon (personal communication, September 15, 2015) has suggested one possible option:

⁴² Bonnefon (2009) does not deny that probabilities are important in reasoning with utility conditionals; he simply argues that it is impossible to integrate probabilities and

RATIONALITY, PRAGMATICS, AND SOURCES

‘If you do not use sunscreen, your risk of skin cancer will not decrease.’

(h 0 A; w - - h)⁴³

Here, not using sunscreen amounts to doing nothing, and therefore has neither utility nor disutility. In contrast, not decreasing one’s risk skin cancer is treated as a loss.

This asymmetry may seem surprising. One way to justify it is to argue for a distinction between the utility of an action and its consequences. There is something more immediate and more certain about the utility of the antecedent: it is, in Bonnefon’s (2009) term, the proximal utility. The utility of the antecedent is the intrinsic, unavoidable utility of an agreeable activity or the intrinsic, unavoidable disutility of a disagreeable activity, in both cases separate from whatever consequences may follow from (not) performing the activity. The utility of the consequent, in contrast, is less immediate and less certain: it is, in Bonnefon’s (2009) term, the distal utility. Given this greater distance and uncertainty, it makes sense to consider these utilities as potential gains and losses. Nevertheless, other utility assignments are plausible, and will be discussed below.

If we adopt the utility assignment above, we can derive a set of predictions for goal frames, which can (and will) be tested experimentally. The goal-framing literature typically explores frames which advocate an action which has at least mild intrinsic disutility but gives rises to consequences with at least mild utility. Table 5.1 reports the aggregate utilities for all possible goal frames based on the above assumptions.

ordinal utilities without making tenuous assumptions. He suggests a possible workaround for this issue.

⁴³ By convention, in a utility grid 0 utility has a target of *A*, which stands for ‘everyone’.

Table 5.1. Utilities for goal frames.

Frame	Antecedent	Consequent	Aggregate
<i>Positive</i>	P1	-	0
	P2	-	+
	P3	--	-
	P4	--	0
<i>Negative</i>	N1	0	-
	N2	0	--
	N3	0	-
	N4	0	--

The aggregate utility corresponds to the strength of the frames as argument. The strength of a positive frame increases as its positive utility increases. A positive frame is persuasive if its aggregate utility is positive: the more positive, the more persuasive. A negative frame is persuasive if its aggregate utility is negative: the more negative, the more persuasive. To elaborate, a negative frame succeeds by being an argument against the antecedent: against not performing the action.

Averaged across the individual frames, positive frames should be weak (on average 0); negative frames should, in contrast, be persuasive (at least -). This higher-level prediction is indistinguishable from the hypothesis of generalized loss aversion, which has already been considered in the literature (Meyerowitz & Chaiken, 1987). At a lower level, there are more distinctive predictions: the strongest arguments are N2 and N4, followed by P2, N1 and N3. Arguments P1 and P4 are unpersuasive; argument P3 is so unpersuasive that it should backfire; that is, it is, in effect, an argument for not doing the antecedent action.

We are now in a position to test the framework above against empirical data.

5.2 Experiment 5.1

Experiment 5.1 is a first attempt at exploring goal framing from the point of view of utility conditionals. The experiment allows a test of two levels of predictions

RATIONALITY, PRAGMATICS, AND SOURCES

based on the utility assignment discussed above: the high level prediction that negative frames should be more persuasive than positive frames; and the lower level predictions about the comparative persuasiveness of individual frames. The experiment assumes Bonnefon's (2009) extended set of simple ordinal utilities: --, -, 0, +, ++. These are realized by the following simple actions and outcomes:

Taking a mildly unpleasant medicine (-)

Undergoing painful surgery (--)

Your risk of catching H1, a minor infection, decreasing (+)

Your risk of catching the DX virus, a life-threatening illness, decreasing (++)

These actions/events were embedded in frames of the following form:

Positive Frame. If you take this mildly unpleasant medicine (undergo this painful surgery), your risk of catching the H1, a minor infection (the DX virus, a life-threatening illness), will decrease.

Negative Frame. If you don't take this mildly unpleasant medicine (undergo this painful surgery), your risk of catching the H1, a minor infection (the DX virus, a life-threatening illness), won't decrease.

In these contexts, the antecedent actions are both preventative behaviours.

Participants saw all the permutations of the utilities shown in Table 5.1, and rated these on a scale of argument convincingness from -5 to 5. This scale was chosen to allow participants to indicate that an argument was so poor that it backfired (an explanation was provided; see below). The experiment was designed to explore the effectiveness of framing using Bonnefon's (personal communication, September 15, 2015) suggested utility assignments. To check whether participants did behave in the suggested way, the experiment included a manipulation check in which participants rated the utility of the actions and consequences.

5.2.1 Methods

Design. This experiment manipulated the full set of utilities from Table 5.1, which correspond to the gain/loss framing. There are two types of frame: positive and negative. The main overarching prediction is that negative frames will be stronger than positive frames. If this prediction holds, further comparisons are possible among the items, to check the precise predictions of the utility grid. This was a within-subjects design.

Participants. 92 participants (42 female; average age 39.30 years) completed the task; 1 participant had previously been excluded since their first language was not English. This experiment used the same system as above for recruiting and remunerating participants. Participants were recruited via the intermediary MTurk Data (www.mturkdata.com). High qualifications were set for the task to improve the quality of the data and maximize the number of native English speakers: participants had to be resident in the US, and have an overall approval rating of 99% for 1,000 previously completed tasks. Participants received a small fee, chosen to exceed the US minimum wage per minute.

Materials. These were eight conditionals: four positive, four negative. The materials manipulate antecedent and consequent utility.

Positive 1: If you take this mildly unpleasant medicine, your risk of catching H1, a minor infection, will substantially decrease. (h – h; w + h)

Positive 2: If you take this mildly unpleasant medicine, your risk of contracting the DX virus, a life-threatening illness, will substantially decrease. (h – h; w ++ h)

RATIONALITY, PRAGMATICS, AND SOURCES

Positive 3: If you undergo this painful surgery, your risk of catching H1, a minor infection, will substantially decrease. (h - - h; w + h)

Positive 4: If you undergo this painful surgery, your risk of contracting the DX virus, a life-threatening illness, will substantially decrease. (h - - h; w + + h)

Negative 1: If you don't take this mildly unpleasant medicine, your risk of catching H1, a minor infection, will not substantially decrease. (h 0 h; w - h)

Negative 2: If you don't take this mildly unpleasant medicine, your risk of contracting the DX virus, a life-threatening illness, will not substantially decrease. (h 0 h; w - - h)

Negative 3: If you don't undergo this painful surgery, your risk of catching H1, a minor infection, won't substantially decrease'. (h 0 h; w - h)

Negative 4: If you don't undergo this painful surgery, your risk of contracting the DX virus, a life-threatening illness, won't substantially decrease. (h 0 h; w - - h)

Procedure. After giving informed consent, participants read the following instructions:

Thank you for taking part in this study. We would like you to imagine that you are talking to a doctor. Your doctor is helping you decide whether to go through some medical procedures. While you are not required or expected to undergo them, there are arguments to consider.

You will be asked to say how convincing you find 8 short arguments. You might read, for example, 'If you take this medicine, you will feel better'. Here, you would be asked to say how convincing this is as an argument for taking the medicine. You will give your answers on a scale from -5 to 5. '-5' means

RATIONALITY, PRAGMATICS, AND SOURCES

that the argument is completely unconvincing. You should understand this as meaning so unconvincing that it backfires. For example, if you give the example argument -5, that would mean that you think the argument above is so bad that it would actually convince you NOT to take the medicine. '5' means that the argument is completely convincing. So if you give the example argument 5, that would mean you think it would definitely convince you to take the medicine.

Please read the arguments carefully. They might sound similar, but the arguments don't repeat. Each one is subtly different. You will also be asked some follow-up questions.

Participants were assigned round-robin style to one of eight surveys containing all eight arguments. In the first of these surveys, the items were ordered according to a random sequence generated using an online random-sequence generator. In the remaining seven surveys, the order was counterbalanced. Participants rated all arguments on the scale provided. After so doing, participants then contemplated the manipulation check. To do so, participants rated the utility of the actions and outcomes mentioned in the arguments. They were instructed as follows: 'Finally, we'd like to know how good or bad you think certain events are. Please reply using the dropdown lists'. The drop-down list comprised the following options always in the following order:

Very bad

Bad

Neither bad nor good

Good

Very Good

RATIONALITY, PRAGMATICS, AND SOURCES

The rated events were the following, the order for which was reversed in half the surveys:

Taking a mildly unpleasant medicine

Not taking a mildly unpleasant medicine

Undergoing painful surgery

Not undergoing painful surgery

Your risk of catching H1, a minor infection, decreasing

Your risk of catching H1, a minor infection, not decreasing

Your risk of catching the DX virus, a life-threatening illness, decreasing

Your risk of catching the DX virus, a life-threatening illness, not decreasing

Finally, participants received debriefing information.

5.2.2 Results

The most fundamental prediction was that negatively framed conditionals would be more persuasive than positively framed conditionals. Figure 5.1 shows the descriptive data.

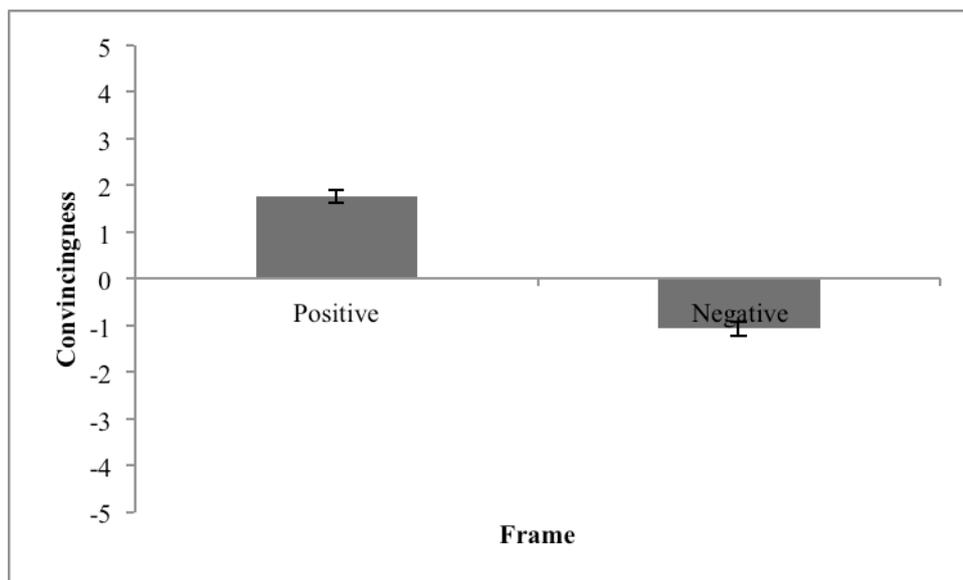


Figure 5.1. Argument convincingness by frame; error bars are standard error

The data clearly suggest the opposite of the predicted trend: positive frames ($M = 1.77$, $SD = 2.78$) are more convincing than negative frames ($M = -1.07$, $SD = 2.99$). While the standard deviations suggest considerable spread, this would be expected because of the variation in the utilities. These data were analysed using a crossed random-effects model in which convincingness was predicted by Frame (positive, negative), and the slope of Frame was allowed to vary randomly across participants and items⁴⁴. The model also included random intercepts for participants and items. These analyses were performed in the Bayes Factor package⁴⁵(Morey & Rouder, 2015). Here, and throughout this chapter, the analysis used uninformative pre-set priors, selecting a wide prior for the fixed effects and a ‘nuisance’ (ultra-wide) prior for the random effects. The latter prior type is recommended for analyses where medium-to-large effect sizes are possible but not of interest, such as variance due to participants (Morey, Rouder, & Jamil, 2015).

A model containing Frame was compared with a null model, which differed only by excluding Frame. There was anecdotal evidence in favour of the model including Frame, $BF = 2.20$. Table 5.2 displays the estimates of the fixed effects. Note that the Bayes Factor package uses effect coding: the intercept is the grand mean; the slopes are deflections from the grand mean. Parameter estimates throughout this chapter are the mean estimate based on 10,000 simulations of

⁴⁴ Note that ‘item’ here, and throughout this chapter, refers to the four combinations of events (taking the medicine and reducing the risk of minor infection; taking the medicine and reducing the risk of the life-threatening illness; and so on), which appear under positive and negative frames.

⁴⁵ Later in the chapter it will be helpful to use Bayesian parameter estimation when exploring the manipulation check data, to minimize the risk of false positives. For consistency, therefore, Bayesian analyses are used throughout this chapter.

RATIONALITY, PRAGMATICS, AND SOURCES

parameter values. The 95% percentile interval reports the middle 95% of these simulations.⁴⁶

Table 5.2. Fixed effects of Frame on Convincingness

Fixed Effect	Parameter	95% PI
Intercept	$b = .35$.16, .55
Positive Frame	$b = 1.14$	-.29, 2.40
Negative Frame	$b = -1.14$	-2.40, .29

The manipulation check gives some insight into why the data do not support the predictions. Table 5.3 shows the frequencies of each category of utility for each antecedent⁴⁷.

Table 5.3. Counts for antecedent utilities

Utility	Medicine	Not Medicine	Surgery	Not Surgery
Very Bad	0	2	49	0
Bad	28	12	30	5
Neither	48	34	8	16
Bad/Good				
Good	14	37	4	24
Very Good	1	6	0	42

These counts suggest that, for ‘Not [Antecedent]’ conditionals, the distributions shift towards more positive ratings. The modal response for taking a mildly unpleasant medicine is neither bad nor good; for not taking it, (narrowly) good. The modal response for undergoing painful surgery is very bad; for not undergoing painful surgery, very good. The manipulation check does not, therefore, support the utility assignments for the antecedents in Table 5.1.

⁴⁶ Together, the mean estimate and 95% PI give a reasonable summary of the data. Note that values outside of the 95% PI can, however, be more probable than values within. Later sections will report the 95% HDI, which is the narrowest interval into which 95% of the estimates fall. This interval can be understood more simply as the 95% most plausible values. This interval is not available in the Bayes Factor package.

⁴⁷ Note that a technical fault led to one participant’s manipulation-check data not being recorded.

RATIONALITY, PRAGMATICS, AND SOURCES

Table 5.4 shows the frequencies of each category of utility for each consequent.

Table 5.4. Counts for consequent utilities; 'Minor Protection' is protection against a minor infection, 'Major Protection' against a life-threatening illness

Utility	Minor Protection	Not Minor	Major Protection	Not Major
Very Bad	1	9	3	43
Bad	7	51	4	36
Neither	7	26	3	9
Bad/Good				
Good	67	5	22	3
Very Good	9	0	59	0

These data show a similar, if more pronounced, pattern to the data on the antecedent utilities. The modal response for being protected against a minor infection is good; for not being protected, bad. The modal response for being protected against a life-threatening illness is very good; for not being protected, very bad. These data are more in line with the utility assignments for the consequent in Table 5.1.

Since participants did not assign utilities in the way predicted in Table 5.1, there is no sense in testing those lower-level predictions. Instead, a supplementary analysis was performed, using participants' actual utility assignments as a predictor. Since participants saw all the manipulation-check questions at once, and the order was only reversed (not randomized) across surveys, this analysis should be viewed as suggestive rather than definitive. To perform this analysis, the utility ratings for negative frames needed transforming: recall that argument strength for positive frames is determined by positive utility; for negative frames, by negative utility. The utilities for the negative frame were transformed by multiplying by -1. Figure 5.2 summarizes the descriptive data.

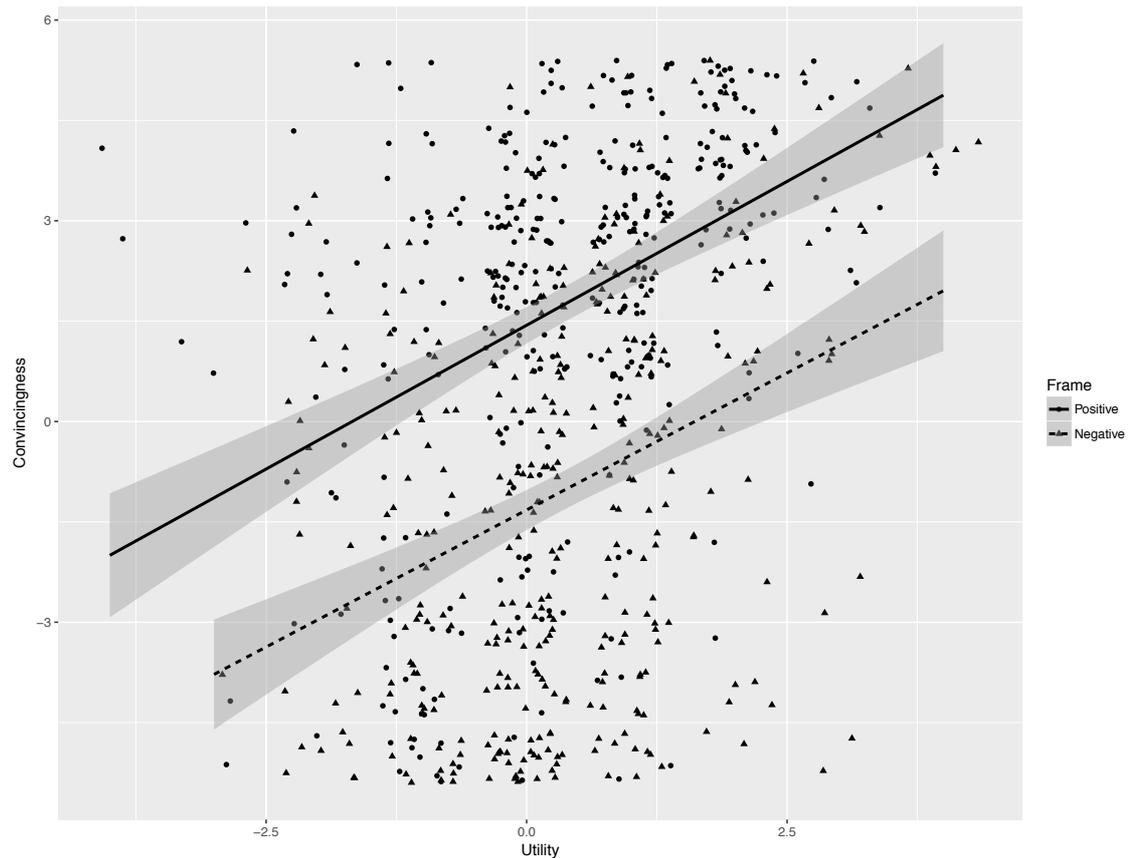


Figure 5.2. Descriptive data by Utility and Frame. Circles are data points for positive frames; triangles are data points for negative frames. The solid line is the regression for positive frames; the dotted line for negative frames. Shading represents the standard error.

The descriptive data suggest a similar effect of Utility for both positive and negative frames. The data were analysed using Bayesian crossed random-effects models. Firstly, a model was fit including Frame and Utility and an interaction term as predictors, with random slopes of Frame across participants and items and random intercepts for participants and items. There was no support for the interaction: compared with the full model, there was moderate evidence in favour of the model without the interaction, $BF_{Interaction} = .21/BF_{Additive} = 4.77$. Table 5.5 reports the fixed effects.

RATIONALITY, PRAGMATICS, AND SOURCES

Table 5.5. Fixed effects of Utility, Frame, and Interaction on Convincingness

Fixed Effect	Parameter	95% PI
Intercept	$b = .36$.18, .54
Positive Frame	$b = 1.13$	-.22, 2.34
Negative Frame	$b = -1.13$	-2.34, .22
Utility	$b = .87$.71, 1.02
Utility * Positive	$b = -.12$	-.30, .06
Utility * Negative	$b = .12$	-.06, .30

Subsequent analyses removed the interaction term, and tested the significance of the remaining fixed effects by comparing the full additive model with a model without the relevant fixed effect (the respective null models). There was very strong evidence for the inclusion of Utility, $BF = 5.18e+23$. There was anecdotal evidence for the inclusion of Frame, $BF = 2.59$. The best model, then, is the full additive model.

Table 5.6 reports the fixed effects for the full additive model.

Table 5.6. Fixed effects of the full additive model

Fixed Effect	Parameter	95% CI
Intercept	$b = .35$.17, .53
Positive Frame	$b = 1.15$	-.11, 2.29
Negative Frame	$b = -1.15$	-2.29, .11
Utility	$b = .87$.71, 1.02

5.2.3 Discussion

This experiment explored the persuasiveness of goal frames from the perspective of the theory of utility conditionals. A plausible utility assignment is that negated antecedents have neutral utility and negated consequents positive utility. According to this assignment, negative frames should be more persuasive. This experiment supports neither an advantage for negative frames nor the assumed utility assignment. The experiment offered evidence that positive frames were, in fact,

RATIONALITY, PRAGMATICS, AND SOURCES

stronger. The manipulation check suggested that this effect was due to negation shifting participants' antecedent utilities towards the positive end of the utility scale, while it shifted participants' consequent utilities towards the negative end of the utility scale. Consequently, negative arguments are rather unconvincing. The supplementary analysis suggested that the frames were considerably less important than participants' actual utility assignments.

There are a number of limitations to the present experiment. Firstly, the manipulation check raises doubts about the suggestive supplementary analysis. All manipulation-check items appeared together, and the order was reversed in half the surveys, not randomized or counterbalanced. It is possible that seeing positive and negative events together prompted an artificial symmetry in utility assignments. Better data would result from separating positive and negative events, and from varying the order more thoroughly. Secondly, the response scale was problematic. A scale from -5 to 5 is not, in itself, problematic, but the description may not have been transparent. The description emphasized that -5 meant that the frame was so unconvincing that it backfired. This description allowed a test of the ultimately redundant lower-level predictions. But is worth replicating these data with a more transparent scale.

5.3 Experiment 5.2

5.3.1 Methods

Design and Materials.

This experiment used the same materials as Experiment 5.1 to explore the effects of framing (Frame) and utility assignments (Utility) on the convincingness of the frames as arguments for undertaking a medical procedure.

Participants.

91 participants (42 female; average age 38.49 years) completed the task. This experiment used the same system as above for recruiting and remunerating participants. Participants were recruited via the intermediary MTurk Data (www.mturkdata.com). High qualifications were set for the task to improve the quality of the data and maximize the number of native English speakers: participants had to be resident in the US, and have an overall approval rating of 99% for 1,000 previously completed tasks. Participants received a small fee, chosen to exceed the US minimum wage per minute.

Procedure.

After giving informed consent, participants read the same instructions as above, except for the description of the scale. The new scale was described as follows:

You will give your answers on a scale from 0 to 10. ‘0’ means that the argument is completely unconvincing. ‘10’ means that the argument is completely convincing.

Participants were assigned round-robin style to surveys. Each survey contained all eight arguments, as in Experiment 5.1. In the first of these surveys, the items were ordered according to a random sequence generated using an online random-sequence

RATIONALITY, PRAGMATICS, AND SOURCES

generator. In the remaining surveys, the order was counterbalanced. Participants rated all arguments on the scale provided. Participants also rated the utilities of two sets of events: (1) the positive antecedents (taking the medicine, undergoing the surgery) and consequents (risk reduction for minor infection, risk reduction for serious illness); and (2) the negative antecedents (not taking the medicine, not undergoing the surgery) and consequents (no risk reduction). Sets (1) and (2) were presented on separate pages, either immediately before or immediately after the arguments, this ordering being counterbalanced across the surveys. Moreover, the order of events within the pages was counterbalanced. Participants were instructed to rate the utility of events on drop-down scales as follows:

‘Firstly, we’d like to know how good or bad you think certain events are.

Please reply using the lists below, where -5 means very bad, 0 means neither bad nor good, and 5 means very good.’

Finally, participants received debriefing information.

5.3.2 Results

Figure 5.3 shows the descriptive data.

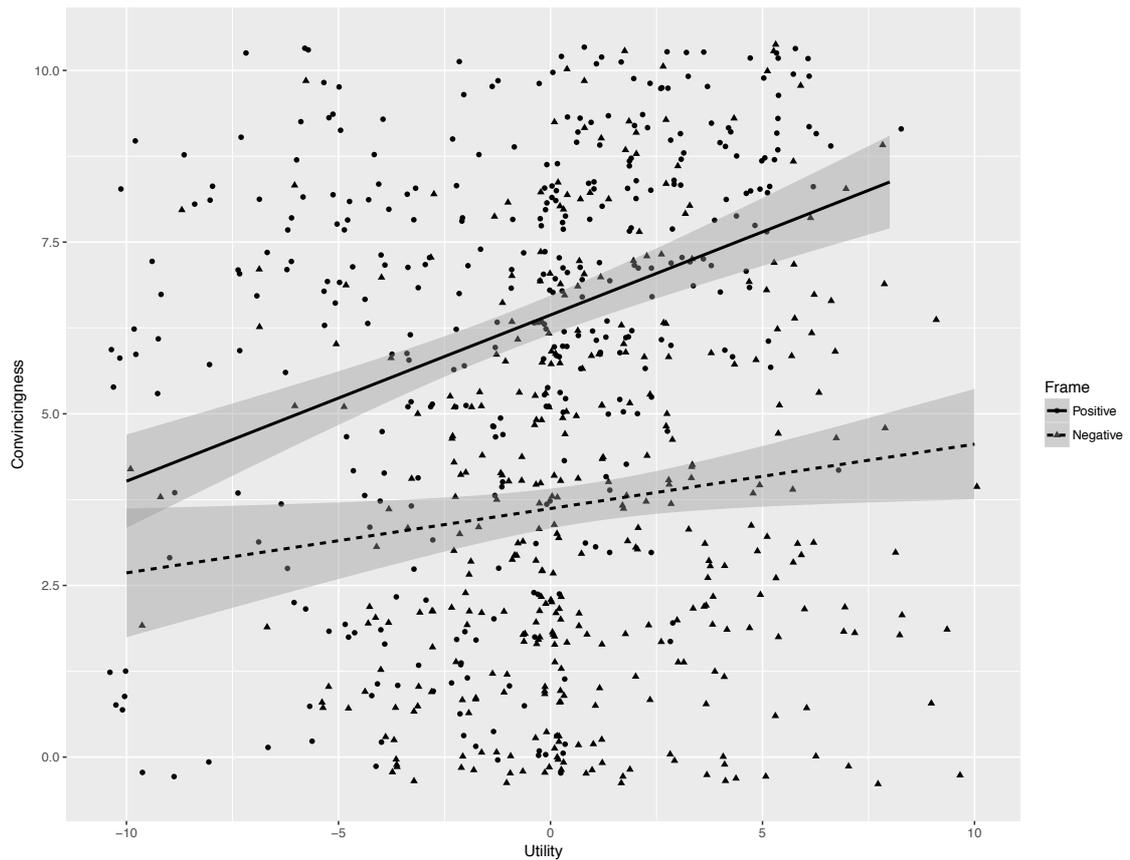


Figure 5.3. Descriptive data by Utility and Frame. Circles are data points for positive frames; triangles are data points for negative frames. The solid line is the regression for positive frames; the dotted line for negative frames. Shading represents the standard error.

These data replicate the basic effects: a linear trend for Utility, with positive frames ($M = 6.20$, $SD = 2.84$) more convincing than negative frames. ($M = 3.70$, $SD = 2.74$).

The data suggest a possible interaction, with a steeper slope for positive frames. As above, these data were analysed using a Bayesian crossed random-effects model in which convincingness was predicted by Frame (positive, negative), Utility and the interaction, and the slope of Frame was allowed to vary randomly across participants and items. The model also included random intercepts for participants and items.

Firstly, a model including the interaction was compared with one including only the main effects. There was anecdotal evidence against including the interaction,

$$BF_{Interaction} = .59, BF_{Additive} = 1.71.$$

RATIONALITY, PRAGMATICS, AND SOURCES

Table 5.7. Fixed effects of Frame, Utility, and Interaction

Fixed Effect	Parameter	95% PI
Intercept	$b = 5.00$	4.80, 5.19
Positive Frame	$b = 1.07$	-.25, 2.27
Negative Frame	$b = -1.07$	-2.26, .25
Utility	$b = .19$.13, .25
Utility * Positive	$b = .05$	-.004, .11
Utility * Negative	$b = -.05$	-.11, .004

The main effects were analysed by comparing the full additive model with models without each main effect. There was anecdotal evidence for the inclusion of Frame, $BF = 2.64$, and very strong evidence for the inclusion of Utility, $BF = 9.46e+8$. The best model, then, is the full additive model, though there was again far stronger evidence for the inclusion of Utility than Frame. Table 5.8 reports the fixed effects for this model.

Table 5.8. Fixed effects for full additive model.

Fixed Effect	Parameter	95% PI
Intercept	$b = 4.95$	4.76, 5.14
Positive Frame	$b = 1.08$	-.23, 2.27
Negative Frame	$b = -1.08$	-2.27, .23
Utility	$b = .20$.14, .26

We turn, now, to the utilities data. These data allow us to explore in greater detail how people assigned utilities to the events mentioned in the arguments. These data were analysed with Bayesian parameter estimation using the BEST package (Meredith & Kruschke, 2013). This method is discussed in more detail in Chapter 6. It suffices, here, to note that the analysis used the default uninformative priors. The method fits a t-distribution to the data, using the following priors: for the mean, the prior is a normal distribution centred on the mean of the data, the spread being

RATIONALITY, PRAGMATICS, AND SOURCES

determined by the precision equivalent to 100 times the standard deviation; for the standard deviation, the prior is a broad uniform distribution from 1/1000 to 1000 times the standard deviation of the data; for the normality parameter, the prior is an exponential distribution giving roughly equal credibility to nearly normal and heavy-tailed distributions (Kruschke, 2013). Since the analysis was exploratory – there were no predictions - we will consider only the posterior distributions of the mean utility ratings for each event. Note that the analyses also yield the 95% HDIs for these distributions: that is, the 95% most plausible values for the mean.

Figure 5.4 contrasts the posterior distributions for the mean utility of taking the medicine and not taking the medicine.

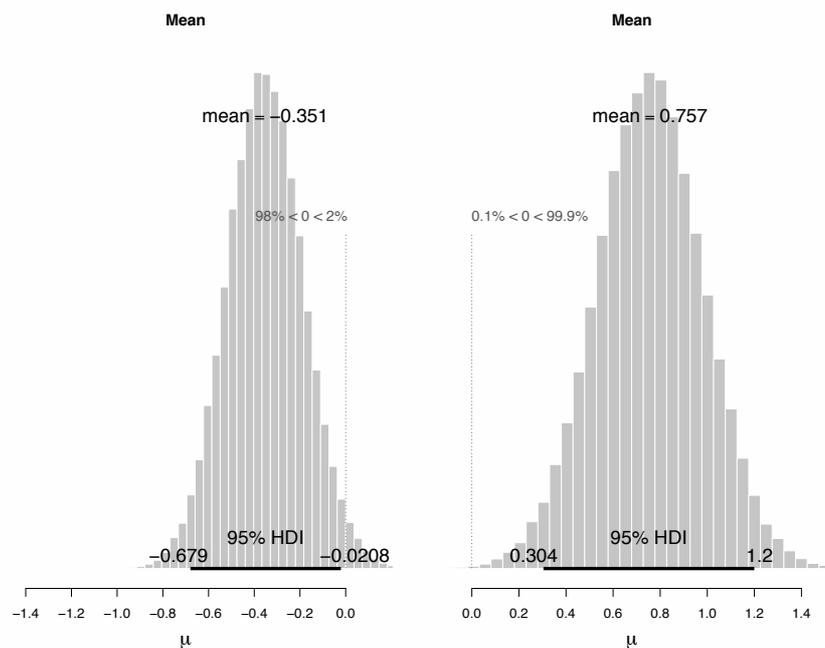


Figure 5.4. Posterior distributions of mean utilities of taking medicine (left) and not taking medicine (right)

Negation flips the utilities: taking the medicine has negative utility ($M_{estimate} = -0.35$) and not taking the medicine has positive utility ($M_{estimate} = .76$). This pattern recalls that in Experiment 5.1, though here taking the medicine has clearer disutility. Figure

RATIONALITY, PRAGMATICS, AND SOURCES

5.5, below, shows the same pattern for undergoing and not undergoing surgery, again similar to that in Experiment 5.1.

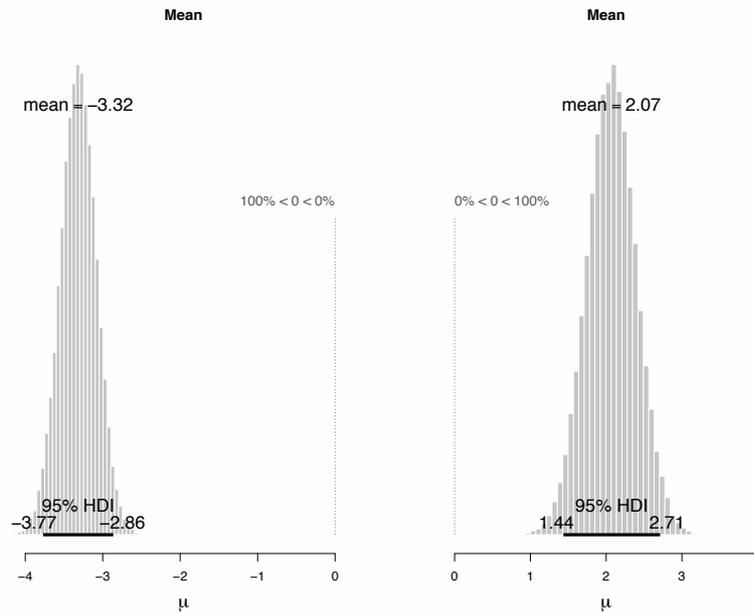


Figure 5.5. Posterior distributions of mean utilities of undergoing surgery (left) and not undergoing surgery (right)

We turn to the utilities of the consequents. Figure 5.6 displays utilities for the minor infection.

RATIONALITY, PRAGMATICS, AND SOURCES

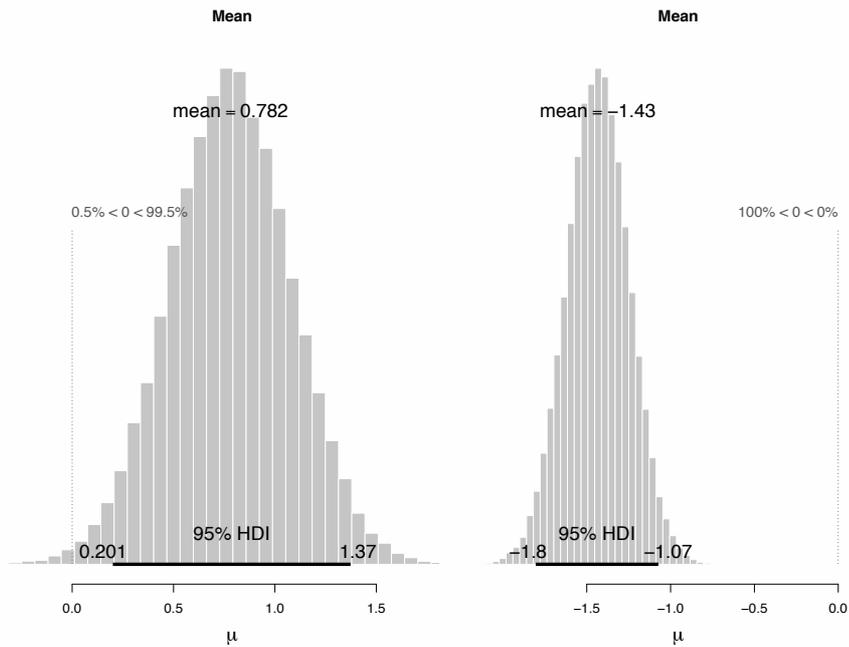


Figure 5.6. Posterior distributions of means of utilities for risk of minor infection decreasing (left) and not decreasing (right)

This figure shows the now characteristic flip of utilities under negation, as do the utilities for the serious illness shown in Figure 5.7.

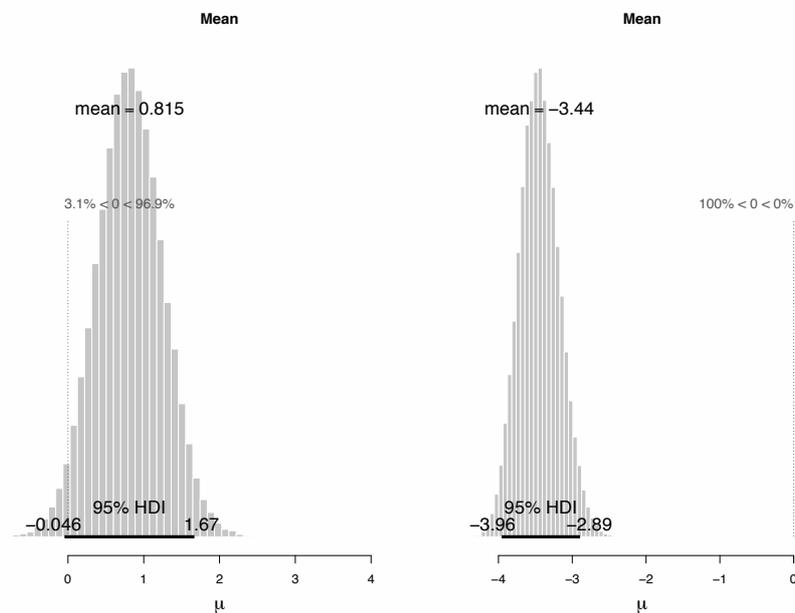


Figure 5.7. Posterior distributions of mean utilities for risk of major illness decreasing (left) and not decreasing risk (right)

It is perhaps surprising, here, that the 95% HDI for decreasing the risk overlaps with 0, albeit by a very small amount, and that there appears to be no discernible difference in utility between the decrease in risk of a minor and major illness.

5.3.3 Discussion

Experiment 5.2 replicated the results of Experiment 5.1 while removing potential confounds. Together the experiments suggest that participants used the utilities to judge the convincingness of the frames as arguments for having a medical procedure. In both experiments, there was some, weaker evidence that, controlling for utilities, there was still an advantage for positive frames. Experiment 5.2 offered greater insight into the utilities: there was a characteristic pattern for negation to flip the utilities.

Why do participants treat the utilities this way? For reasons that will become clear later on, I will largely restrict my attention to the utility of the antecedent. The effect of negation recalls the shifting reference point of Prospect Theory (Tversky & Kahneman, 1981). While Prospect Theory does not strictly apply to this experiment – for the same reasons as cited in the introduction to this chapter – considering the similarities will prove informative. Negated clauses were evaluated, not relative to the actual world, but relative to the hypothetical world in which the action being considered was, in fact, taken. Consider the conditional ‘If you don’t undergo this painful surgery, your risk of catching the DX virus, a life-threatening illness, won’t decrease.’ The antecedent utility is evaluated relative to a world in which the hearer undergoes painful surgery; the consequent utility is evaluated relative to a world in which the hearer’s risk of catching the DX virus decreases. In these worlds, the antecedent utility is a gain; the consequent utility is a loss (though see Experiment 5.3 for qualifications on the consequent).

We return, at this point, to the topic of pragmatics; for pragmatics has been suggested as a mechanism for the shifting reference point. Although we cannot call on Prospect Theory to explain the present data, we can call on pragmatics. McKenzie and colleagues (McKenzie & Nelson, 2003; Sher & McKenzie, 2006) have treated this issue from the perspective of attribute framing, the simple type of framing when, for instance, participants respond better to a product described as 95% fat-free than one described as having 5% fat. McKenzie and colleagues argued that such attribute frames imply an increase from a reference point in the relevant attribute: our 95% fat-free product is more fat-free than a reference point, the average of that type of product on the market. They have supported this account with experimental data suggesting that people are sensitive to reference points. When participants read that a glass was half-full, they inferred that it had previously been empty; when they read that a glass was half-empty, they inferred that it had previously been full. Likewise, when describing a glass, participants used ‘half-full’ if it had been previously empty and ‘half-empty’ if it had previously been full. McKenzie and colleagues suggest that reference points are shifted through natural-language pragmatics.

Such research is highly suggestive. Although the account is underspecified from the point of view of the linguistic processing, it suggests an explanation for the current data. Indeed, with negation, things may be somewhat clearer. Negative sentences show interesting patterns of assertability. For instance, the sentence ‘A whale isn’t a fish’ seems reasonable, but ‘A whale isn’t a bird’ is decidedly strange. We can reasonably expect that some people might think a whale is a fish - after all, it looks and acts like a fish in various ways - but not that people might think a whale is a bird (Wason, 1965). Negative sentences, in other words, invite the supposition that what is negated is expected to have been the case: the sentence ‘John isn’t home’

invites the supposition that John was expected to be at home (Clark, 1974).⁴⁸

Psycholinguistic research has explored this phenomenon with negative natural-language quantifiers, such as ‘few’, and suggests that negation pragmatically invites a focus on the complement set (Moxey, 2006; Moxey & Sanford, 1986, 2000; Paterson, Sanford, Moxey, & Dawydiak, 1998). To apply this account to the present topic, it makes little sense for you to consider assigning utility to not having an unpleasant procedure unless there is an expectation that you will have the procedure. Negation communicates precisely this expectation.

Negative sentences bring to mind the complement set: the alternative scenarios in which the corresponding affirmative sentence holds. This is part of a more general phenomenon whereby utterances bring to mind sets of alternatives. This phenomenon may be implicated in framing more generally (Geurts, 2013). Geurts presents two sentences based on the materials of the Asian disease paradigm. Imagine a scenario in which 600 lives are at stake.

It is good that 200 people survived.

It is good that 400 people died.

Let us assume, furthermore, that the numbers have an exact reading: *exactly* 200 people survived, and *exactly* 400 people died. Although these sentences refer to a mathematically equivalent situation, they appear to contradict each other. Geurts suggests that the contradiction results from the alternatives that are brought to mind. On his account, in such contexts, evaluative adjectives like ‘good’ bring to mind

⁴⁸ Clark uses the term ‘supposition’ to hedge against conflating two potentially distinct phenomena. As he notes, sentences such as ‘Stop cheating on your exam’ and ‘Don’t stop cheating on your exam’ both presuppose that the addressee is cheating on his/her exam, whereas the sentences ‘John is present’ and ‘John isn’t present’ do not both suppose that John was expected to be present. Nevertheless, we can say that the speaker of a negative sentence presupposes that the corresponding affirmative sentence was expected to be true.

ordered scales of alternatives: it is good that 200 people survived; it would have been better if 201 people had survived; and so on. The sentence ‘It is good that 400 people died’ is odd because it gives rise to the following scale: it is good that 400 people died; it would have been better if 401 people had died; and so on (though adding ‘only’ before the number term appears to reverse this scale). The implied scales contradict each other; hence the apparent contradiction between the two sentences. It is, as Geurts argues, a short step to framing: simply drop the explicit evaluative language and have participants themselves perform the evaluation, either explicitly on a rating scale or implicitly through their choice of options. Indeed, Geurts offers explanations of both attribute and Asian-disease framing, though these have yet to be empirically tested.

The preceding paragraphs make an important point for the thesis: natural-language pragmatics has a profound effect on the assignment of utilities. When this point is considered alongside the theory of utility conditionals, however, a tension arises which needs resolving. In the utility-conditionals theory, it is utilities that guide pragmatic inference, not pragmatic inference that guides utility assignment. A potential solution is to consider different types of pragmatic phenomena. As we have seen, a sentence such as ‘He’s not fat’ invites the supposition that the addressee might think otherwise. This supposition is rather basic and hard to deny. It is assumed to be part of the common ground and is not part of the ‘at issue content’ (on at issue content, see, e.g., Tonhauser, 2012). Hence, the addressee cannot simply reply ‘That’s not true’; doing so would amount to saying that the ‘he’ in question is, indeed, fat. The appropriate response is, for example, to say ‘I never thought that he *was* fat’. Something rather different is at stake in utility conditionals. Take a sentence such as ‘If I let you get away with it, then my boss will fire me’. Here, the

utilities guide us to the implicit – but very much at issue - point, of the utterance: the speaker means ‘I cannot let you get away with it’. The picture, then, is of a complex interplay between pragmatics and utilities: low-level pragmatic phenomena feed into the assessment of utilities, and utilities feed back into the generation of higher-level pragmatics. This interplay has not previously featured in the theory of utility conditionals, in large part because negation has also not featured in the theory.

Although we have already seen that Prospect Theory does not apply to the current experiment, we might nevertheless be tempted to invoke the prospect-theory-inspired predictions of Rothman and Salovey (1997). After all, the frames are about prevention behaviours and, as Rothman and Salovey predict for these behaviours, the positive frame is more persuasive. It is worth underlining why the data do not, in fact, support the Rothman and Salovey account. As the introduction to this chapter showed, the Rothman and Salovey predictions rely on prevention behaviours being viewed as low risk because they have certain, positive outcomes. Indeed, in this experiment the consequences of undertaking the behaviour are all positive. However, this experiment manipulated the utility of both antecedent and consequent; the aggregate utility significantly predicted the convincingness of the frames. There is no space in the Rothman and Salovey account for this predictor.

Although Experiment 5.2 offers a substantial improvement on Experiment 5.1, there are still important limitations. Firstly, there is an ambiguity over the wording of the consequent. Take, for instance, ‘your risk of contracting the DX virus, a life-threatening illness, decreasing’. This consequent could correspond equally well to the risk increasing and the risk staying the same, two events which would intuitively have rather different utilities. This ambiguity will need addressing before we can judge whether the foregoing discussion of utilities applies equally to the

antecedent and consequent. Lastly, there was anecdotal evidence that the type of frame predicted argument convincingness even after controlling for the utilities.

There is no predictor in this design to account for this effect. One possibility is the conditional probability. Although participants could have read these conditionals as deterministic, they may nevertheless have read them as probabilistic. Since participants were asked to judge the convincingness of each item as an argument for performing the antecedent action, they may have been particularly focused on the conditional probability $P(\text{PositiveConsequent} \mid \text{PositiveAntecedent})$. The positive frames likely provide better evidence for this probability being high than the negative frames.

5.4 Experiment 5.3

The previous section identified an ambiguity in the wording of the consequent. This study addressed the ambiguity by adapting the test on the utilities. The experiment simply presented the consequent events from Experiments 5.1 and 5.2 and required participants to judge their utility on a scale from -5 to 5. The experiment was between-participants. The consequent events were described in three different ways: the risk of an illness increasing; the risk not changing; and the risk staying the same. The last two phrasings are different ways of presenting the same event: the risk remaining at a constant value. However, given the account of negation in discussing the previous experiment, it seems plausible that the phrasings could lead to different utility assignments. Accordingly, we predict main effects of severity and phrasing, but no interaction. Additionally, we will ask whether the ‘stay the same’ and ‘not change’ phrasings have disutility or are neutral.

5.4.1 Methods

Design.

The design manipulated two independent variables: event severity (a minor infection, a life-threatening illness), and phrasing (the risk increasing, not changing, staying the same). The experiment was between-participants.

Materials.

The items were as follows:

Your risk of catching H1, a minor infection, increasing

Your risk of catching H1, a minor infection, not changing

Your risk of catching H1, a minor infection, staying the same

Your risk of contracting the DX virus, a life-threatening illness, increasing

Your risk of contracting the DX virus, a life-threatening illness, not changing

Your risk of contracting the DX virus, a life-threatening illness, staying the same

Participants.

241 participants (102 female; average age 34.71 years) completed the task; the data for 5 other participants were excluded since these participants reported a first language other than English. This experiment used the same system as above for recruiting and remunerating participants. Participants were recruited via the intermediary MTurk Data (www.mturkdata.com). High qualifications were set for the task to improve the quality of the data and maximize the number of native English speakers: participants had to be resident in the US, and have an overall approval rating of 99% for 1,000 previously completed tasks. Participants received a small fee, chosen to exceed the US minimum wage per minute.

Procedure.

After giving informed consent, participants read the same instructions:

Thank you for taking part in this study. We would like you to imagine that you are talking to a doctor. Your doctor is helping you decide whether to go through some medical procedures. While you are not required or expected to undergo these procedures, there are arguments to consider.

On the following page, you'll be told about an event. You will be asked how good or bad this event is, on a scale from -5 to 5, where -5 is very bad, 0 is neither good nor bad, and 5 is very good.

Participants were then assigned, round-robin style, to a survey where they gave demographic information and rated a single item. Finally, they received debriefing information.

5.4.2 Results

Figure 5.8 shows the descriptive data by condition.

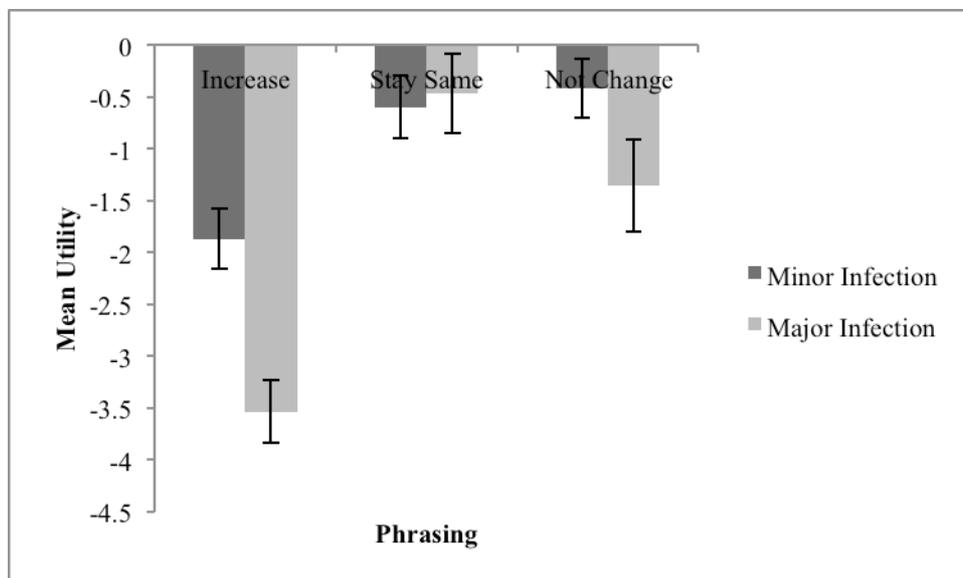


Figure 5.8. Mean utilities by condition; error bars are standard error

There is a hint, here, of an interaction. For the minor infection, the utilities decrease steadily from the 'increase' condition ($M = -1.87$, $SD = 1.79$) to the 'stay same' condition ($M = -.6$, $SD = .192$) and the 'not change' condition ($M = -0.41$, $SD =$

RATIONALITY, PRAGMATICS, AND SOURCES

1.84). For the major infection, the utilities decrease from the ‘increase’ condition ($M = -3.54$, $SD = 1.94$) to the ‘stay same’ condition ($M = -0.46$, $SD = 0.30$) before increasing again to the ‘not change condition ($M = -1.36$, $SD = 0.44$). Note that the ‘stay same’ condition shows a different pattern from the other conditions, with the minor infection rated as having (very) slightly higher disutility than the major infection. The data were analysed, as above, using Bayesian regression analyses using a wide prior for the fixed effects. Two models were compared, with and without interactions: there was anecdotal evidence against including the interaction, $BF_{Interaction} = .88/BF_{Additive} = 1.14$. Table 5.9 reports the parameter estimates for the full model.

RATIONALITY, PRAGMATICS, AND SOURCES

Table 5.9. Parameters for model with interaction

Fixed Effect	Parameter	95% PI
Intercept	$b = -1.41$	-1.67, -1.15
Minor Infection	$b = .40$.15, .66
Major Infection	$b = -.40$	-.66, -.15
Increase	$b = -1.25$	-1.61, -.89
Stay Same	$b = .80$.43, 1.16
Not Change	$b = .46$.09, .81
Interaction: Minor & Increase	$b = .38$.03, .74
Interaction: Minor & Stay Same	$b = -.39$	-.74, -.03
Interaction: Minor & Not Change	$b = .001$	-.34, .35
Interaction: Major & Increase	$b = -.38$	-.74, -.03
Interaction: Major & Stay Same	$b = .39$.03, .74
Interaction: Major & Not Change	$b = -.001$	-.35, .34

There was good evidence for the inclusion of Severity, $BF = 8.24$, and very strong evidence for the inclusion of Phrasing, $BF = 29964006$. We now perform pairwise comparisons on Phrasing using independent-samples Bayesian t-tests in the Bayes Factor package, again using a wide prior for the effect. There was extremely strong evidence that the ‘increase’ condition had stronger disutility than the ‘stay the same’ condition, $BF = 17273783$, and ‘not change’ conditions, $BF = 83312.19$. There was moderate evidence that the ‘stay the same’ and ‘not change’ conditions did not differ, $BF_{Difference} = .19/BF_{Null} = 5.26$.

Finally, it was predicted that phrasing ‘stay same’ and ‘not change’ would still be reliably negative. These predictions were tested with one-sample Bayesian t-tests. For the ‘stay same’ condition, there was equivocal evidence for a difference, $BF = 1.24$); for the ‘not change’ condition, in contrast, there was strong evidence for disutility, $BF = 39.52$, though note the moderate evidence above for equal means.

Since Bayesian analyses are not sensitive to sampling intentions⁴⁹ (Kruschke, 2013, 2015), we can also directly compare Experiment 5.2’s ambiguous ‘not decrease’ conditions with the corresponding conditions in Experiment 5.3. These comparisons take the form of Bayesian independent-samples t-tests. Given the ambiguity of the 5.2 materials, we should not expect an exact match, because participants may have differed in their interpretations, but these analyses should nevertheless be suggestive. Because the designs of the studies differed, the simplest analysis is not possible, that is, directly comparing the ambiguous ‘not decrease’ items as a whole with the three separate phrasing conditions. Recall that, in Experiment 5.2, all participants rated all utilities; hence the data are not independent. We can, nevertheless, compare the individual conditions (‘not decrease’ plus minor infection, ‘not decrease’ plus major infection) with the new phrasings. For the major-infection data, there was moderate evidence for equal means between the ‘not decrease’ and ‘increase’ conditions ($BF_{Difference} = .29/BF_{Null} = 3.42$); there was strong evidence for a difference between the ‘not decrease’ and ‘stay the same’ ($BF = 76249.17$) and the ‘not decrease’ and ‘not change’ ($BF = 41.36$) conditions. For the minor-infection data, there was again moderate evidence for no difference between

⁴⁹ Frequentist t-tests would arguably be inappropriate here, because p-values depend on researchers’ sampling intentions (Kruschke, 2013, 2015). Being drawn from different

the ‘not decrease’ and ‘increase’ conditions ($BF_{Difference} = .33$, $BF_{Null} = 3.03$). There was anecdotal evidence of a difference between the ‘not decrease’ and ‘stay the same’ conditions ($BF = 2.29$), and good evidence of a difference between the ‘not decrease’ and ‘not change’ conditions ($BF = 7.28$). Taken together, these data suggest that the closest correspondence is between the ‘not decrease’ and ‘increase’ conditions.

5.4.3 Discussion

This simple experiment sheds light on the way participants interpreted the negative conditionals in Experiments 5.1 and 5.2. In particular, the consequent ‘your risk of [illness] will not decrease’ is ambiguous between the risk not changing and the risk increasing. This experiment addressed the ambiguity using the wording ‘your risk of [illness] will increase’ and two variants of no change: ‘your risk of [illness] will stay the same’ and ‘your risk of [illness] will not change’. Unsurprisingly, the ‘increase’ wording led to clearly greater disutility than the other two wordings, which were not reliably different from each other. Both the ‘increase’ and ‘not change’ wordings resulted in reliable disutility; the ‘stay the same’ wording did not.

This experiment suggests that, in Experiments 5.1 and 5.2, negative frames depended for their persuasiveness on the ‘increase’ reading of the consequent. Accordingly, if negative frames were expressed unambiguously as the ‘stay the same’ or ‘not change’ readings, then they would fall even further behind positive frames. This point emphasizes further that we cannot simply assume the equivalence

experiments, the data for these analyses were not collected with a specific sample in mind. Bayesian analyses do not suffer this disadvantage.

of positive and negative frames. There is no obviously correct reading of ‘not decrease’, and the information content depends on how this phrase is interpreted. Since readings equivalent to ‘stay the same’ and ‘not change’ are widespread in the framing literature, the current findings raise questions about the convincingness of these negative frames and how the convincingness bears on the equivocal findings of the literature.

The data also prompt a caveat to the discussion of the previous results. In discussing Experiment 5.2, we considered whether utility assignments depend on pragmatics: more specifically, whether negations invite the hearer to suppose that the negative event was expected. This suggestion helps to explain why participants assigned positive utility to not doing an unpleasant action. Superficially, the same account applied to the consequents. But the data above suggest that this is not straightforwardly the case. For, if negation is the crucial factor, then ought there not to be a reliable difference between the ‘stay the same’ and ‘not change’ conditions?

The lack of a difference between the relevant conditions prompts a more nuanced view of negation and utilities. In the present experiments, the antecedents are distinct actions: the action is either taken or not taken. In such cases, negation may give rise to a clear and distinct expectation: it was expected that the hearer would take the action. The consequents, however, are less clear. Experiment 5.3 sought to remove the ambiguity with the wordings ‘increase’, ‘stay the same’ and ‘not change’, but ‘not change’ is, in fact, still ambiguous in a crucial respect. Change may be expected, but the direction of change is unclear. In context, it seems plausible that participants understood ‘not change’ as implying that a decrease in risk was expected, because they were instructed to imagine that they were talking to a doctor about undergoing procedures to protect themselves against possible illness.

RATIONALITY, PRAGMATICS, AND SOURCES

We do not, however, have direct evidence that this was how people interpreted ‘not change’. These points suggest that more research is needed on the interplay between negation, expectations and utility assignment. Experiments could, for instance, cue different expectations and then assess the convincingness of different frames.

The data above serve as a reminder that, to understand framing effects, we must understand how participants interpret the frames. Throughout this chapter, we have seen that negation, expectations, and utilities are key factors. To illustrate, let us return to a pair of frames from (Apanovitch et al., 2003):

Positive frame: If you decide to get HIV tested, you may feel the peace of mind that comes with knowing about your health.

Negative frame: If you do not decide to get HIV tested, you may not feel the peace of mind that comes with knowing about your health.

Experiments 5.1 and 5.2 seem to imply that the antecedent in the positive frame will receive disutility and the antecedent in the negative frame will receive utility. At first blush, the consequents may seem equivalent: the use of ‘may’ implies a lack of certainty, from which it trivially follows that ‘may not’ is also true. But the experimental data above prompt us to look more closely. The consequent describes an event, feeling peace of mind, that is not ambiguous in the way that ‘change’ is. Plausibly, then, the negation-and-expectation account will hold for this consequent: that is, participants will infer that the expected outcome is feeling peace of mind. As a result, the consequents may differ in the expectations they give rise to.

In sum, the data in this experiment argue for close consideration of individual frames. Subtle differences can, in principle, give rise to differences in convincingness. These differences will obtain within, as well as across, types of

frame; it seems plausible, then, that the effectiveness of frames will differ widely according to context.

5.5 General Discussion

This chapter has considered the relationship between utility and pragmatics, using goal framing as a test case. The experiments show how people interpret positive and negative frames, and how these interpretations feed into judgments of convincingness. The chapter set out to test predictions from the theory of utility conditionals, assuming a particular utility assignment. These predictions were not supported. However, there was good evidence that utilities matter to the convincingness of goal frames. Positive frames were more convincing than negative frames, the advantage being explicable in terms of the utilities of antecedents and consequents. Negative frames had antecedents with high enough utility to offset the negative utility of the consequent. The overall negative frames, then, had little utility or disutility, and consequently tended to be weak arguments. These effects held across two different ways of measuring utility and convincingness. Finally, Experiment 5.3 probed further, unpicking the ambiguous consequent of the negative frames. The data suggested that participants had interpreted the lack of a risk reduction as an increase in risk. Alternative wordings had far less disutility ('not change') or no reliable disutility ('stay the same'), suggesting that negative frames with these wordings would be still less persuasive.

The experiments above emphasize the link between pragmatics and utilities; hence, the link between pragmatics and rationality. We have seen that the utility assignments can be understood with a negation-and-expectation account. We originally supposed that not performing an action has zero (dis-)utility, since it amounts to doing nothing. However, negation implies that the negated event was

expected to happen. If this expectation is strong, then it makes sense for people to treat missing out on an expected positive as a loss and escaping an expected negative as a gain. Experiment 5.3 complicated this picture by suggesting that it was the ambiguity of the consequent, and not expectations, that drove the utility assignments. In particular, the ‘not change’ phrasing did not differ reliably from ‘stay the same’. Nevertheless, the data do not contradict the negation-and-expectation account, because the ‘not change’ phrasing does not set up a clear expectation about the direction of change. The materials were under-specified in this respect. We should expect ‘not change’ to have different utilities if, say, the context suggested that the hearer should expect an increase in risk – say, they were moving to an environment in which an illness was endemic – or that the hearer should expect a decrease in risk – say, they were taking up a new healthy lifestyle. The data overall, then, are still compatible with a complex relationship between pragmatics and utilities: negations and expectations play a key role in determining utilities, which then feed back into justifying higher level pragmatic inferences. Given that people routinely make decisions based on descriptions of actions and outcomes, the literature would benefit from a far broader treatment of how people judge utilities under verbal negation.

Utilities do not explain all the variance in the data. There is anecdotal evidence, from Experiments 5.1 and 5.2, that positive frames have an advantage over negative frames even controlling for utilities. It is possible that the remaining variance is due, at least in part, to probabilities. Although we have largely ignored probabilities in this chapter, we have already seen that probabilities play an important role in learning from conditionals and in the assessment of arguments. Indeed, probabilities are also implicit in the negation-and-expectation account: an expectation can, presumably, be understood as a probabilistic belief about the future.

The present experiments do not allow a straightforward test of probabilities, because the materials were designed to focus on utilities and, hence, suggested a deterministic relationship between antecedent and consequent. While it is possible that people nevertheless understood the conditional probabilistically, testing this possibility would require asking participants to provide a probability for a risk. It is not clear that this is a meaningful quantity for participants. Future experiments could, however, explore the role of probabilities by switching to more clearly probabilistic relationships, and could make it easier to estimate these probabilities by using real-world materials.

So far, the discussion has, in effect, been addressing the premise-interpretation component of Crupi's model. Knowing the information content of frames is a prerequisite for judging the rationality of people's reasoning. While it is premature to judge this rationality – contrary, for instance, to nudge-inspired accounts – we can preview how to assess rationality at a later date. To judge rationality, in this case, the crucial quantity seems to be the expected utility of the conditional. So, for a conditional 'If p then q':

$$EU(\text{Conditional}) = U(p) + P(q|p) * U(q)$$

(Evans et al., 2008, p. 114)

More descriptive accounts argue that we should expect people to judge only the specific circumstances mentioned (Evans et al., 2008). By this reasoning we should expect sensitivity to the expected utility above alone. However, a more standard decision-theoretic approach would suggest that people should judge the expected utilities of all the actions under consideration and select an action by optimizing some utility function (Peterson, 2009). By this reasoning we should expect sensitivity to the expected utility of both the positive frame and the negative frame. It

remains to be seen whether people show sensitivity to either of these quantities.

Nevertheless, if we wish to judge people's rationality in reasoning from frames, we should aim to measure their behaviour against these yardsticks.

The discussion above argues for a larger research program, investigating how people assign utilities under verbal negation, whether people are sensitive to probabilities, and whether they ultimately judge the expected utility of the frames. This program will directly bear on the framing literature. There is, however, a mismatch between the framing literature and the current studies. The current studies used convincingness ratings, a dependent measure which has been shown to be reliable in existing research (Hahn & Oaksford, 2007). The framing literature, on the other hand, prefers measures such as attitude scales, ratings of intention to perform, or (more rarely) actual compliance with recommendations, with effects differing across these measures (Gallagher & Updegraff, 2012). It will be important to test whether the suggested research program produces results which generalize across different dependent measures⁵⁰.

This chapter has focused heavily on the relationship between pragmatics and rationality. But it is easy to see the importance of sources. Previous chapters have argued that, whenever a conditional appears in an experiment, it can be considered a testimonial conditional: the conditional has a source even if that source remains implicit. In the present experiments, the source was held constant: participants were asked to imagine that the source was their doctor. In real life, people, of course, receive information from sources of differing reliabilities; as Chapters 2 to 4 showed,

judgements of reliability impact on various probabilities associated with the conditional. We should expect, then, that source reliability will affect belief change from utility conditionals and, hence, will also matter to convincingness judgements and perhaps also to ratings of attitudes or intentions to perform or, indeed, to compliance behaviours. But there is no reason, *a priori*, to restrict belief change to the probability. If people hear a reliable source saying a utility conditional, they may also change their own utility assessments. In effect, a reliable source saying that an outcome is desirable may lead us to value that outcome more. These points further illustrate the tight connections between the components of the triangular scheme.

⁵⁰Generalizability could be tested in another respect: namely, how utility is measured. The decision-making literature does not widely use ratings scales to assess utility. A future study could test whether the effects survive using alternative utility assessments such as the Becker-DeGroot-Marshak method (Becker, Degroot, & Marschak, 1964).

6 Testimony and Source Reliability ⁵¹

This thesis has focused on a triangular scheme for the relationship between rationality, pragmatics, and sources. The preceding four chapters have explored different parts of this scheme. Chapters 2 to 4 focused on the relationship between sources and pragmatics, using the case study of conditionals in simple testimonial contexts. Chapter 5 focused on the relationship between rationality and pragmatics, using the case study of goal framing. This chapter will focus on the relationship between rationality and sources, using the case study of testimony. More specifically, the chapter will discuss the normative issues that are key to learning from sources; it will introduce two alternative normative models; and it will use these models to explore people's behaviour. The chapter returns to the domain of testimony but from a different perspective. This contrast leads to the introduction of a crucial distinction for this thesis - between interpreting and accepting – so that pragmatics will resurface once again.

6.1.1 Introducing Testimony

Testimony, in the philosophical sense, is part of epistemology. Philosophers of testimony consider how we should, and how we do, learn from other people's statements: when, in other words, should we trust people's claims? These very same concerns animate research in a wide range of disciplines, albeit rarely under the label 'testimony'. Witness, for instance, classic research on the 'wisdom of the crowds' (e.g. Galton, 1907; Surowiecki, 2005), on judgment and decision making (e.g. Birnbaum & Mellers, 1983; Birnbaum & Stegner, 1979; Birnbaum, Wong, & Wong,

1976), on persuasion (e.g. Petty & Cacioppo, 1984, 1986; Petty, Cacioppo, & Goldman, 1981); and burgeoning research in developmental psychology (for a review, see Mills, 2013), in computer science (for a review, see Artz & Gil, 2007), as well as in philosophy (e.g. Bovens & Hartmann, 2003; Coady, 1992).

Testimony is an issue of fundamental social relevance. For we ‘massively depend on communication with others’ (Sperber et al., 2010) - on the information we receive from other people – and communication itself is testimonial, requiring a ‘stance of trust’ (Sperber et al., 2010). That is, when interpreting other people’s utterances, we must be prepared to change our own beliefs. This stance of trust contrasts with Davidson’s principle of interpretive charity, whereby we supposedly take someone to be communicating ‘a set of beliefs largely consistent and true by our own standard’ (Davidson, 1984, p. 137; cited by Sperber et al., 2010). To illustrate, take Sperber et al.’s (2010, p. 368) example. We are asked to imagine that Barbara has asked Joan to bring a bottle of champagne to a dinner party. The following dialogue occurs:

Andy (to Barbara): A bottle of champagne? But champagne is expensive!

Barbara: Joan has money.

We are asked to imagine, further, that Andy has previously assumed that Joan is an underpaid academic. On Sperber et al.’s (2010) account, if Andy applies the principle of interpretive charity, he interprets ‘Joan has money’ to mean ‘some money, as opposed to none’ or, perhaps, an amount consistent with the salary of an underpaid academic. But if Andy is prepared to trust Barbara and revise his beliefs

⁵¹ This chapter reports is adapted from Collins, Hahn, von Gerber, and Olsson (2015, 2017).

about Joan, he can recover the intended meaning, ‘enough money to be easily able to afford champagne’.

Although we can gain great benefits from information transmitted between people, we risk trusting inexperienced sources, and we are vulnerable to being misled or deceived (Sperber et al., 2010). In other words, testimony is problematic from the point of accuracy and honesty. People can be decidedly unreliable sources even when well intentioned, by being inaccurate: take, for example, forgetting and the reconstructive nature of memory (for discussion, see Hahn, Oaksford, & Harris, 2012; on reconstructive memory, see, e.g., Loftus, 1975). People can also be decidedly unreliable sources by being dishonest. There is evidence, for instance, that people systematically engage in small acts of deception, exploiting opportunities to cheat as long as they can do so without challenging their self-conception, and even deceiving themselves in the process (see, e.g., Chance, Norton, Gino, & Ariely, 2011; Mazar, Amir, & Ariely, 2008).

Such systematic unreliability mandates vigilance towards people as sources and towards the content of their messages: it mandates, in Sperber et al.’s (2010) term, ‘epistemic vigilance’. An epistemically vigilant person may adopt an initial stance of trust towards a source, but will ultimately distinguish between comprehending a message and accepting it (Sperber et al., 2010). It is tempting, considering classic research, to dismiss out of hand the existence of epistemically vigilant sources. As we saw in Chapter 1, Gilbert and colleagues (D. T. Gilbert et al., 1990, 1993) famously argued that people tend towards blind credulity. In their paradigm, participants read sentences such as ‘A Monishna is a star’ which were followed by the label ‘true’ or ‘false’. When disruptions were applied to false trials, participants tended to recall the information as true, suggesting a default assumption

RATIONALITY, PRAGMATICS, AND SOURCES

that new information is true. However, this paradigm arguably confounds negation and informativeness: sentences such as ‘A Monishna is not a star’ are radically uninformative; a ‘Monishna’ could be infinitely many things (Hasson et al., 2005). Once this confound is removed, there is no difference in memory for false and true trials (Hasson et al., 2005).

There is further promise of epistemic vigilance in the developmental literature (for a review, see Mills (2013). For example, children as young as 14 months old show some ability to prefer reliable, and disprefer unreliable, informants (Poulin-Dubois, Brooker, & Polonia, 2011; Poulin-Dubois & Chow, 2009). Children as young as 3 years old show sensitivity to verbally expressed uncertainty (Sabbagh & Baldwin, 2001); to expertise (Robinson, Champion, & Mitchell, 1999); and to past accuracy (Ganea, Koenig, & Millett, 2011). At around 6 years old, children show sensitivity to deception (Mascaro & Sperber, 2009); at around 8 years old, to a source’s self-interest (Mills & Keil, 2005); and at around 10 or 11 years old, to a source’s partiality (Mills & Keil, 2008) and to distortions due to a source’s self-report of evaluative traits (Heyman & Legare, 2005).

There is considerable evidence that sensitivity to sources and message content continues into adulthood. This evidence comes from the psychology of persuasion (for a recent review, see Petty & Briñol, 2008). The early persuasion literature treated source and message as separate components. For Kelman (1958), people accepted a claim either by accepting an argument for it, a process he called internalization, or by identifying with the source, a process he called identification (for discussion, see Petty & Briñol, 2008). More recently, the literature has been dominated by more complex, dual-route models: most famously, the Elaboration Likelihood Model (‘ELM’; e.g. Petty & Cacioppo, 1984; Petty & Cacioppo, 1986;

RATIONALITY, PRAGMATICS, AND SOURCES

Petty, Cacioppo, & Goldman, 1981) and the Heuristic-Systematic Model (e.g. Chaiken, Liberman & Eagly, 1989). According to these models, there are two routes for persuasion: a central route, in which the audience focuses on systematically analyzing arguments; and a peripheral route, in which the audience focuses on heuristics such as general impressions or surface features of an argument. The routes also differ in how extensively audiences think about the issue in hand: in the technical parlance, elaboration. In the central route, there is high elaboration: audiences use the central route when they are able, or motivated, to analyze an issue in depth. In the peripheral route, there is low elaboration: audiences use the peripheral route when they are less able or motivated.

In its current incarnation, the ELM identifies five ways in which sources can induce persuasion (Briñol & Petty, 2009; Petty & Briñol, 2008). (1) Under conditions of low elaboration, sources can act as simple, heuristic cues. A classic study presented students with arguments of various strengths for a new university exam, and manipulated both source reliability and personal relevance (Petty, Cacioppo, & Goldman, 1981). When personal relevance was low, persuasiveness was due to source reliability; when personal relevance was high, persuasiveness was due to argument strength, that is, the actual content of the persuasive message. High or low personal involvement thus influences the choice of routes to persuasion. (2) Under conditions of high elaboration, sources can act as an argument or evidence. When an attractive source testifies to the effectiveness of a beauty product, the source's appearance is visual evidence for the effectiveness of the product (Petty & Briñol, 2008). In other words, although source characteristics generally feature in the peripheral route, where they provide a heuristic cue, source information can, on occasion, have evidential value on the central route. (3) Sources can affect

RATIONALITY, PRAGMATICS, AND SOURCES

metacognition. For example, it has been found that when source information comes after an argument, credible sources influence not the valence of thoughts, but rather people's confidence in their thoughts (Briñol, Petty, & Tormala, 2004). (4) Sources can bias thinking. For example, source expertise can affect the direction of thoughts, so long as the message is ambiguous and the task is important (Chaiken & Maheswaran, 1994). (5) Sources can affect the extent of thinking. For example, when there are multiple sources for a claim, people tend to think longer, magnifying differences attributable to argument strengths: strong arguments become more persuasive; weak arguments, less persuasive (Harkins & Petty, 1981). In other words, (4) and (5) allow source information to affect analytic processing in ways that go beyond evidential value, by moderating the direction and amount of analytic thinking that takes place.

The persuasion literature clearly suggests that people are sensitive to information about sources and, as such, it is relevant to the question of testimony. But testimony is a more basic phenomenon, whose proper treatment is logically prior to a full theory of persuasion. The persuasion literature focuses on persuasive texts in which informal arguments are presented⁵². But these texts are composed of claims and reasons which are not just abstract propositions, but also testimonial claims put forward by an arguer (Hahn, Oaksford & Harris, 2012). Moreover, a proper treatment of testimony is a step towards computational models of how individual claims and reasons are processed, which in turn are a step towards constrained, predictive models of persuasion. Such models could help to transform the field of

⁵² Osgood & Tannenbaum (1955) developed an account of attitude change from testimonial claims, but this account was driven by the attitude, in the social-psychological

persuasion, in which there is a proliferation of rich descriptive accounts but a conspicuous lack of detailed processing accounts.

Models of testimony have already proved a rich source of inspiration for psychological studies. We have already seen one such Bayesian model in Chapter 4 and how this model can be extended to capture data on testimonial conditionals. At the core of these Bayesian models are three principles: the reliability of the source is evaluated; the probability of the claim is evaluated; and these evaluations are combined in some way. Bayesian modelling of testimony has inspired the development of norms for informal argumentation and the assessment of human behaviour against these norms (Hahn et al., 2009; Hahn, Harris, & Oaksford, 2013; Hahn et al., 2012). It has also inspired research in developmental psychology (Shafto, Eaves, Navarro, & Perfors, 2012), legal testimony (Fenton, Neil, & Lagnado, 2013; Friedman, 1987; Schum, 1981), and consensus among climate scientists (Hahn et al., 2016).

Since Bayesian models of testimony are normative and provide readily testable predictions, they allow research to contribute directly to the Great Rationality Debate in cognitive science. It is worth dwelling, for a moment, on their normative basis. Bayesianism in general can be justified on the grounds that, under certain conditions, Bayesian belief revision is demonstrably optimal (Leitgeb & Pettigrew, 2010b, 2010; Rosenkrantz, 1992). Bayesian models of testimony typically represent probabilities as subjective degrees of belief; revising these subjective probabilities in a Bayesian way maximises coherence. The case for Bayesian models,

sense, towards the source and claim, and not by a critical examination of them as in testimony.

in particular, is especially strong when a phenomenon is clearly inferential (Hahn, 2014).

The Bayesian approach to testimony has generated two striking models which the experiments below will explore. In both of these models, source reliability plays a central role: a claim fares better if its source is considered reliable. But the models do not assume that sources come with their reliability obviously labelled. It is, of course, quite plausible that people will make initial judgments of trustworthiness based on stereotypes, personality characteristics, or mannerisms and behaviours such as voice characteristics, gestures or eye movements (DePaulo et al., 2003; Vrij, Granhag, & Porter, 2010). Such presumed evidence can contribute to the setting of prior reliability. Important, too, is the content of what someone says. Source reliability should obviously be damaged by a catastrophic error or an intentional lie. But what about when the truth is not so obvious? Both models allow the plausibility of claims to influence source-reliability judgments. There is, then, a bidirectional relationship between the content of a claim and the reliability of the source.

We have already seen one of these Bayesian models of testimony: Bovens and Hartmann's (2003) simple Bayesian belief network, discussed in Chapter 4. It will be useful to briefly recapitulate the details of this model. The model comprises three random variables: the variable *Hyp* is a proposition about the world, and has the states true and false; the variable *Rel* is the reliability of the speaker, and has the states true (reliable) and false (unreliable); the variable *Rep* is the report that '*Hyp* is true', and has the states true (the report is made) and false (the report is not made). The variable *Rep* is dependent on *Rel* and *Hyp*, but these variables are (marginally) independent of each other. When a source is reliable ($Rel = \text{True}$) it makes a report;

RATIONALITY, PRAGMATICS, AND SOURCES

when it is unreliable ($Rel = False$) it makes a report with some random degree of probability. If a report is made, then Rel and Hyp become dependent (i.e. are conditionally dependent), and their values are simultaneously revised. The revision depends on the prior probabilities of Hyp and Rel . If $P(Hyp = True) < .5$, then reliability ($P(Rel=True)$) will be revised downwards; if $P(Hyp = True) > .5$, then reliability will be revised upwards. If the reliability is zero, then there will be no change in belief. For any higher reliability, $P(Hyp = True)$ will increase, although this increase may be very small indeed with very low reliabilities; the higher the reliability, the bigger the revision in $P(Hyp = True)$.

An alternative model has been developed by Olsson and Angere, reported, for instance, in (Olsson & Vallinder, 2013). This model has important commonalities and differences. For present purposes, the most important are these. Both models predict that an expected claim ($P(Hyp=True) > .5$) will increase reliability and that an unexpected claim ($P(Hyp=True) < .5$) will decrease it. However, the Olsson and Angere model takes a different approach to reliability: specifically, while reliable sources cause an increase in $P(Hyp=True)$, unreliable sources cause a decrease. In other words, sources can be, not just unreliable, but anti-reliable: their testimony can be negatively correlated with the truth.

Anti-reliability is superficially similar to so-called backfire effects. In these effects, people decrease their belief in a claim despite its assertion (for discussion, see Cook & Lewandowsky, 2011). These effects are, however, importantly different. Backfire effects can occur when someone tries to correct popular misinformation: citing myths and debunking them can lead to a short-term correction but longer-term strengthening of the myth, as the audience finds the myth even more familiar and forgets the details of the correction (Skurnik, Yoon, Park, & Schwarz, 2005).

RATIONALITY, PRAGMATICS, AND SOURCES

Backfire effects can also occur when someone argues against a belief by providing too many arguments, overburdening the audience (Cook & Lewandowsky, 2011; Schwarz, Sanna, Skurnik, & Yoon, 2007). Lastly, backfire effects can occur when someone tries to persuade an audience that has an established world-view. Here, different audiences can polarize, despite receiving the same information, and strengthen their pre-existing views (Nyhan & Reifler, 2010; Taber & Lodge, 2006). These effects thus rely on the presence of strong pre-existing views and on specific processing mechanisms. In contrast, anti-reliability applies to any assertion and relies only perceptions of the source.

Together, the Bayesian models make useful predictions about how people revise their beliefs in simple testimonial contexts. These predictions will form the basis of the experiments below. The models predict that people will use their beliefs about the world and about a source in revising their belief. Both models predict that reliability will increase in response to expected claims and decrease in response to unexpected claims. Both models predict that belief in a claim will increase in response to reliable sources; the models differ in their predictions for belief change from unreliable sources. If we draw on the persuasion and developmental literatures, these predictions seem plausible; but there is a lack of clear evidence on simple testimonial contexts. And it is these testimonial contexts that are closest to the crucial notion of epistemic vigilance. If people treat testimony critically, in the way the models predict, then there is a clear difference between comprehension and acceptance.

The chapter reports three experiments which explore the model predictions in simple testimonial contexts. Experiment 6.1 examined belief change from more or less reliable sources. Experiment 6.2. replicated this study. Experiment 6.3 examined

change in reliability after expected and unexpected claims Experiment 6.4 replicated this study. Finally, Experiment 6.5 used a different methodology which avoided any reference to source reliability in simple contexts with multiple sources, to test whether participants spontaneously used message content to revise beliefs about a source. It is worth noting that the experiments were not intended to provide a definitive test of competing models. Indeed, the two models can, to some extent, be viewed as complementary: Harris et al. (2016), for instance, implemented a Bayesian network which combines elements of both models. Rather, the experiments were intended as existence proofs: as providing evidence for the set of behaviours above and, in particular, source anti-reliability.

6.2 Experiment 6.1: Belief change

6.2.1 Method

Predictions. This experiment tested the following predictions. Firstly, both models predict that reliable sources should increase belief in a claim. Secondly, the Bovens and Hartmann (2003) model predicts that unreliable sources should have, at worst, no effect on belief change. Thirdly, the Olsson and Angere model predicts that unreliable sources can decrease belief in a claim.

Design. This experiment followed a 2x2 between-subjects design with the following factors: Claim Expectedness (Expected, Unexpected) and Source Reliability (High, Low).⁵³

⁵³ The original design aimed to test a somewhat more complex set of predictions than discussed here; hence the 2 x 2 design. However, these predictions required lower scores on prior ratings than were actually achieved. The one-way predictions are appropriate to the achieved values.

Participants. 91 people (45 women; average age 38.62) completed online surveys posted on a US-hosted website for academic research (<http://psych.hanover.edu/research/exponnet.html>). Participants were recruited both through this website and through university e-mail lists.

Materials & Procedure. Participants read brief texts about six topics⁵⁴. Each text took the following form. Participants first read a claim and rated its convincingness by responding to the question ‘How convincing is the claim?’ on a Likert-style scale from 0 (not at all convincing) to 10 (completely convincing). For example,

‘One of the best remedies against a severe cough is valium’.

Participants were then presented with a source making this claim:

Now imagine that Michael, who is a clinical nurse specialist, told you the following: ‘One of the best remedies against a severe cough is valium’.

Following this, participants re-rated the convincingness of the claim on the same Likert- scale. The full set of items are as follows:

Unexpected and expected claims; expected claim in brackets
(Unreliable and reliable sources; unreliable sources in brackets)

⁵⁴ The materials and design for Experiments 6.1 and 6.3 were developed by Professors Ulrike Hahn and Erik Olsson, and Dr Ylva von Gerber.

RATIONALITY, PRAGMATICS, AND SOURCES

1. Claim: One of the best remedies against severe cough is valium (lots to drink, hot or cold).

[Convincingness Rating]

Now imagine that Michael, who is a clinical nurse specialist (drug addict), told you the following: ‘One of the best remedies against severe cough is valium (lots to drink, hot or cold).’

[Rating]

2. Claim: The temperature in the Kitchen Mate oven varies a lot (keeps very steady), which is perfect for a fluffy and crispy bread.

[Rating]

Now imagine that Paula, who is a baker (a designer kitchen Kitchen Mate salesperson working on commission), told you the following: ‘The temperature in the Kitchen Mate oven varies a lot (keeps very steady), which is perfect for a fluffy and crispy bread.’

[Rating]

3. Claim: The Australian horse ‘Thunderbolt’, who has beaten the British horse ‘Lightening’ in the majority of races entered this season, will lose to (beat) ‘Lightening’ at the upcoming Cheltenham Festival.

[Rating]

Now imagine that Robert, who is a senior sports reporter and has predicted the winner in the last 10 races he covered (a junior sports reporter and has failed to predict...), told you: ‘The Australian horse ‘Thunderbolt’, who has beaten the

RATIONALITY, PRAGMATICS, AND SOURCES

British horse ‘Lightening’ in the majority of races entered this season, will lose to (beat) ‘Lightening’ at the upcoming Cheltenham Festival.’

4. Claim: In 2013, the maximum temperature in Stockholm in June was 15 degrees (23 degrees)

[Rating]

Now imagine that Mary, who is a retired meteorologist (who is five years old and was given a weather station for Christmas), told you: ‘In 2013, the maximum temperature in Stockholm in June was 15 degrees (23 degrees).’

[Rating]

5. Claim: The Land Rover Discovery has no problems (has problems) with the electricity and very

cheap spare parts.

[Rating]

Now imagine that Patrick, who is a car enthusiast (a used car salesperson), told you: ‘The Land Rover

Discovery has no problems (has problems) with the electricity and very cheap spare parts.’

[Rating]

6. Claim: The Eclipse nightclub in Detroit (Ibiza) has the reputation of one of the coolest nightclubs in the world.

[Rating]

Now imagine that Emma, who is a 26-year-old DJ and has established herself, has a leading figure in the club scene (who is a 45 year old housewife with three young children (between 5 and 12) who enjoys knitting)

and regularly attends the Eclipse nightclub with her friends, told you:

‘The Eclipse nightclub in Detroit (Ibiza) has the reputation of one of the coolest nightclubs in the world.’

[Rating]

6.2.2 Results & Discussion

All data in this chapter were analysed using robust Bayesian parameter estimation. The analyses are, in effect, Bayesian equivalents of classical one-sample t-tests (for Experiments 6.1 to 6.4) and independent-sample t-tests⁵⁵ (for Experiment 6.5). The Bayesian analyses are useful because they provide richer information than the classical tests – a posterior distributions over parameter values – and are not dependent either on assumptions about the data (e.g. normality) or on sampling intentions (Kruschke, 2013). The Bayesian analyses are also invaluable when testing models, because the analyses can lead to both rejection and acceptance of the null hypothesis (Kruschke, 2013).

Firstly, change scores were calculated by subtracting the initial rating from the posterior rating; these were then averaged across items to create a mean change score for each participant. The data were then analysed in the following way, following Kruschke (2013). The analyses do not assume that the data are normally

⁵⁵ Recall that these experiments were intended as existence proofs. The experiments do not test whether one of the Bayesian models provides a general account of testimony. As such, the aim is not to generalize beyond the current set of materials. To so generalize would require a different design (avoiding one-sample tests), to permit the crossed random-effects analyses used in other chapters. On the generalization issue, see Baayen, Davidson, and Bates, (2008); H. H. Clark (1973); Raaijmakers, Schrijnemakers, and Gremmen (1999).

RATIONALITY, PRAGMATICS, AND SOURCES

distributed, but instead describe the data with a t-distribution, which allows heavy tails. T-distributions have three parameters: the mean, μ ; standard deviation, σ ; and normality, ν . Where the value of the normality parameter is large (ca. 100), the distribution is nearly normal; where it small, the distribution is heavy tailed (Kruschke, 2013).

The one-group analysis for this experiment estimates the most credible parameter values, given the data, for the following model:

$$P(\mu, \sigma, \nu | D) = \frac{P(D | \mu, \sigma, \nu) * P(\mu, \sigma, \nu)}{P(D)}$$

The denominator is approximated using Markov Chain Monte Carlo (MCMC) methods, which simulate thousands of combinations of parameter values (for more technical details, see Kruschke, 2013). The analyses were carried out in R (R Core Team, 2015) and JAGS using the packages BEST (Meredith & Kruschke, 2013) and rjags (Plummer, 2003). The default values of the BEST programs were appropriate for the current design. By default, the MCMC chain has 100,000 steps, with no thinning to correct for autocorrelation. The default priors are uninformative. Since this is the first study to explore these content/source predictions, uninformative priors are justified. The prior for μ is a normal distribution centred on the mean of the data, the spread being determined by the precision equivalent to 100 times the standard deviation; for σ it is a broad uniform distribution from 1/1000 to 1000 times the standard deviation of the data; for ν it is an exponential distribution giving roughly equal credibility to nearly normal and heavy-tailed distributions (for further details, see Kruschke, 2013).

The remainder of this section reprises the predictions, and reports the corresponding posteriors for the parameters. To decide whether the parameter

RATIONALITY, PRAGMATICS, AND SOURCES

estimates (dis-)confirm the predictions, we need two further concepts: the highest density interval (HDI) and the regional of practical equivalence (ROPE). The HDI spans the most credible (highest probability) values of the posterior distribution: for instance, a 95% HDI, which we will use throughout, covers 95% of the distribution, and the values within it have a total probability of .95 (Kruschke 2013, 2015). When assessing predictions, we can ask whether the 95% HDI includes a specific point value: for example, for the null hypothesis, zero. In reality, requiring a point value may be too stringent. In such cases, a ROPE can prove helpful: values within this region are considered practically equivalent to the comparison value. Kruschke (2015) recommends that, in the absence of clear guidelines in the field, researchers establish a regional of practical equivalence (ROPE) around the comparison value, from -0.1 to 0.1 . Below, we will apply the ROPE to effect sizes, so that the relevant comparison value will be zero with a ROPE from -0.1 to $+0.1$. We will base our evaluations of the experimental predictions on these effect sizes and corresponding ROPEs. If the 95% HDI falls entirely outside of the ROPE, there is a clear effect; if it falls entirely within the ROPE, there is a null effect. In this case, the 95% HDI for effect size falls outside this conventional ROPE. Where there is overlap, the data do not allow a clear decision for the specific HDI and ROPE. It may, nevertheless, be informative to consider how much overlap there is, as this will give some indication of weaker conclusions.

- (1) Reliable sources should increase belief in a claim
- (2) (i) Unreliable sources should decrease belief in a claim. OR
(ii) Unreliable sources should not affect belief in a claim.

Figure 6.1 displays the mean belief change by reliability.

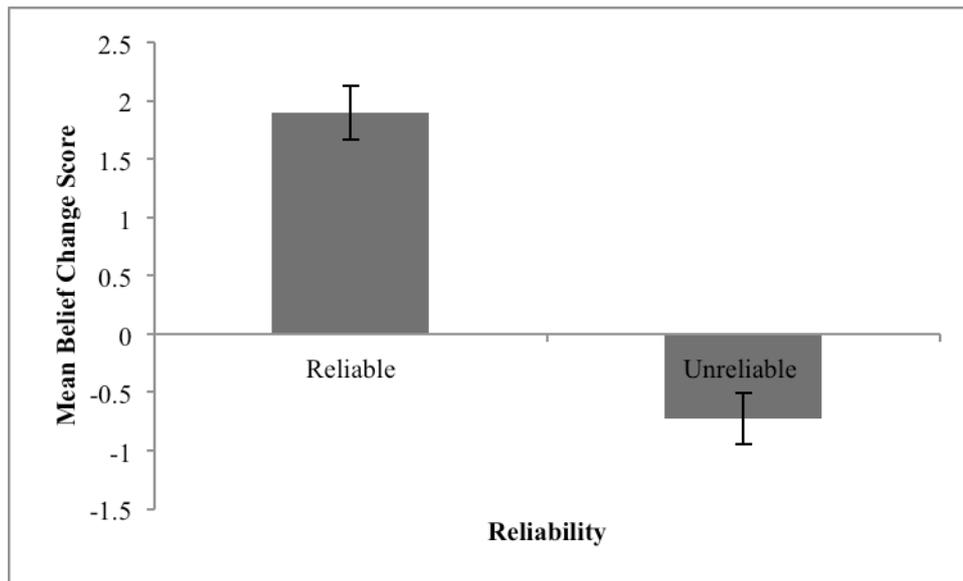


Figure 6.1. Mean belief change by reliability; error bars are standard error

These means show the predicted increase in belief in the claim in response to testimonial evidence from a reliable source, and a decrease in response to the same evidence when coming from an unreliable source. In other words, the data suggestive anti-reliability (2i). These findings were analysed with two one-group analyses with a comparison value of 0, analogous to classical one-sample t-tests.

Reliable Sources. The mean estimate for μ was 1.84 (95% HDI [1.37, 2.3]). The modal estimate for σ 1.39 (95% HDI [1.01,1.81]. The modal estimate for $\log_{10}(v)$ was 1.37 (95% HDI [.42, 2,04]. Lastly, the modal estimate for effect size – $(\mu-0)/\sigma$ – was 1.31 (95% HDI [.87, 1.80]), which falls outside the conventional ROPE. Figure 6.2 shows the posterior distribution for effect size and the ROPE. This analysis, then, shows that reliable sources credibly increased belief in a claim.

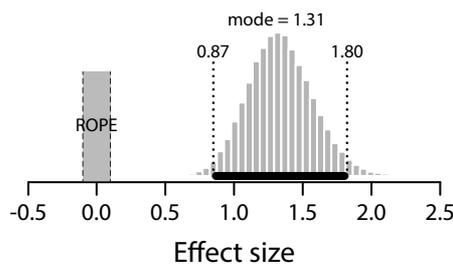


Figure 6.2. Posterior distribution of effect size of belief change from reliable sources. ROPE from -0.1, to 0.1; dotted lines are 95% HDI

Unreliable Sources. The mean estimate for μ was $-.72$ (95% HDI $[-1.15, -.29]$). The modal estimate for σ was 1.46 (95% HDI $[.97, 1.88]$). The modal estimate for v was 1.36 (95% HDI $[.37, 1.99]$). Lastly, the modal estimate for effect size was $-.49$ (95% HDI $[-.87, -.17]$), which falls outside the conventional ROPE. Figure 6.3 shows the posterior distribution for effect size and the ROPE. Thus, this analysis shows that unreliable sources credibly decreased belief in a claim.

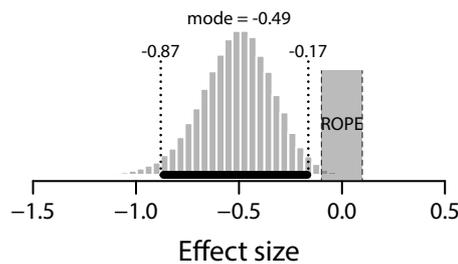


Figure 6.3. Posterior distribution of effect size of belief change from unreliable sources. ROPE from $-.1$ to $.1$; black bar depicts 95% HDI

Summary. These data therefore support both predictions (1) and (2)(i): reliable sources increased belief in a claim; unreliable sources decreased belief in a claim. The data offer support, then, for source anti-reliability. The novelty of the findings argues for replication. Moreover, the data from Experiment 6.1 were collected via a university-hosted website for online experimental studies; this sample of self-selecting, interested volunteers is likely biased toward university students and staff. Experiment 6.2 was posted on Amazon Mechanical Turk. Although samples on Mechanical Turk are also not representative of the general population, they are considered more diverse than college samples (Paolacci & Chandler, 2014), and, most importantly, are likely to be different in composition than the sample of Experiment 6.2. This offers a useful further test of the effects.

6.3 Experiment 6.2: Replicating belief change

6.3.1 Methods

Design. The experiments followed the same design as Experiments 6.1: a 2x2 between-subjects design with the following factors: Claim Expectedness (Expected, Unexpected) and Source Reliability (High, Low).

Materials and Procedure. Experiment 6.2 used the same materials and procedure as Experiment 6.1. Two minor changes were made to the materials to adapt them for a predominantly US audience. Firstly, for the Stockholm item, temperatures were given in both Centigrade and Fahrenheit. Secondly, for the nightclub item, US locations were given: Manhattan as an expected location for a prestigious nightclub, and Des Moines, Iowa, as an unexpected location. The Range Rover item was also removed because, in Experiment 6.1, participants' prior beliefs showed that the intended expectedness manipulation had not worked.

Participants. 79 people (27 women; average age 33.38) completed online surveys posted on Amazon's Mechanical Turk as a small job (HIT). The HIT was posted by an intermediary, MTurk Data. Participants were rewarded with a small fee equivalent to \$0.20 per minute, calculated to exceed the rate of the US minimum wage. To maximize engagement and maximize the number of native English speakers, high qualifications were posted. To complete the task, participants needed to be resident in the US, Canada or UK, have a 99% approval rating for their previous HITs, and to have completed 1,000 approved HITs. In addition, a qualification guaranteed that people could not participate if they had completed the previous experiment. One participant's data (not included in the above count) was excluded because that participant reported a first language other than English.

6.3.2 Results & Discussion

- (1) Reliable sources should increase belief in a claim.
- (2) (i) Unreliable sources should decrease belief in a claim. OR
(ii) Unreliable sources should not affect belief in a claim.

Figure 6.4 shows the descriptive data for Experiment 6.2. Qualitatively, the same patterns are observed as in Experiment 6.1.

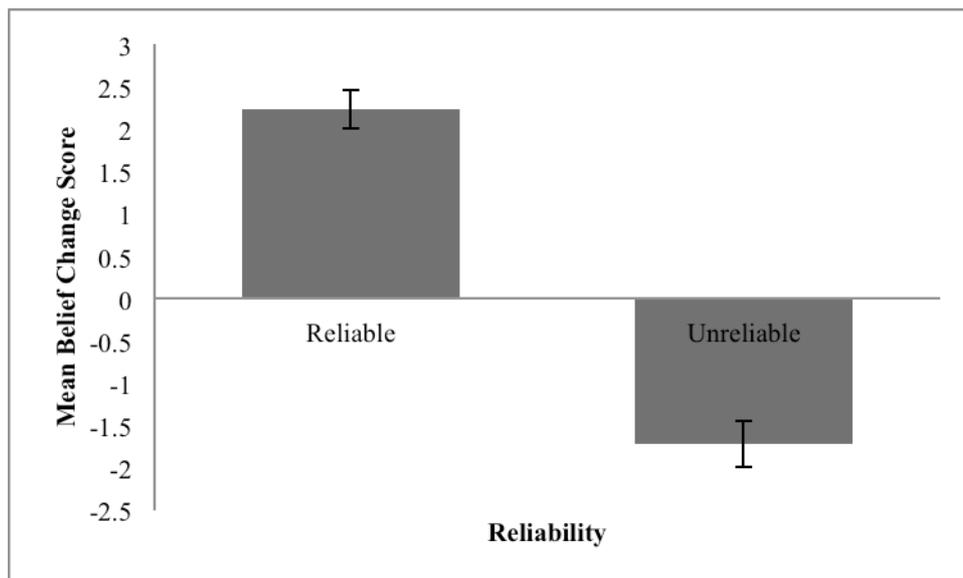


Figure 6.4. Mean belief change by reliability; error bars are standard error

Reliable Sources. The mean estimate for μ was 2.23 (95% HDI [1.76, 2.7]). The modal estimate for σ 1.39 (95% HDI [1.09, 1.79]). The modal estimate for $\log_{10}(v)$ was 1.49 (95% HDI [.75, 2.1]). Lastly, the modal estimate for effect size – $(\mu-0)/\sigma$ – was 1.58 (95% HDI [1.08, 2.10]), which falls outside the conventional ROPE. Figure 6.5 shows the posterior distribution of the effect size and ROPE. This analysis, then, shows that reliable sources credibly increased belief in a claim.

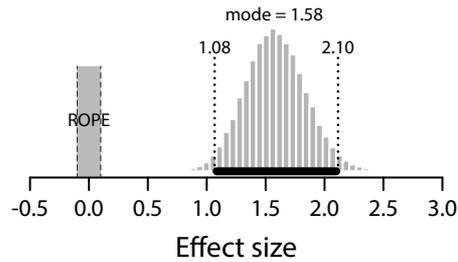


Figure 6.5. Posterior distribution of effect size of belief change from reliable sources. ROPE from -0.1 to 0.1 ; black bar depicts 95% HDI

Unreliable Sources. The mean estimate for μ was -1.73 (95% HDI $[-2.28, -1.18]$). The modal estimate for σ was 1.58 (95% HDI $[1.21, 2.11]$). The modal estimate for ν was 1.48 (95% HDI $[.6, 2.07]$). Lastly, the modal estimate for effect size was -1.08 (95% HDI $[-1.53, -.63]$), which falls outside the conventional ROPE. Figure 6.6 shows the posterior distribution of effect size and the ROPE. This analysis shows that unreliable sources credibly decreased belief in a claim.

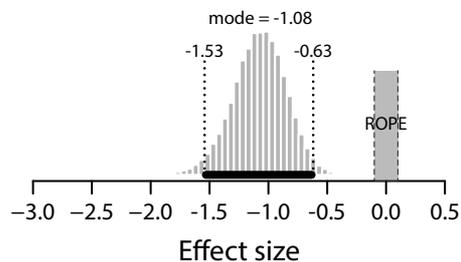


Figure 6.6. Posterior distribution of effect size of belief change from unreliable sources. ROPE from -0.1 to 0.1 ; black bar depicts 95% HDI

Summary. These data therefore support both predictions (1) and (2i): reliable sources increased belief in a claim; unreliable sources decreased belief in a claim. The data replicate the effects in Experiment 6.1, providing further support for source anti-reliability.

6.4 Experiment 6.3: Reliability change

6.4.1 Methods

Design. This experiment followed a 2x2 between-subjects design with the following factors: Claim Expectedness (Expected, Unexpected) and Source Reliability (High, Low).

Participants. 131 people (45 women; average age 39.83) completed online surveys posted on a US-hosted website for academic research (<http://psych.hanover.edu/research/exponnet.html>). Participants were recruited both through this website and through university e-mail lists.

Materials & Procedure. Participants read texts on the same six topics as in Experiments 6.1 and 6.2. The only difference concerned the dependent variables. Instead of providing an initial judgment on the convincingness of the claim, the participants first read about the source and rated its reliability by responding to the question ‘How reliable do you think [source name] is?’ on a Likert-style scale from 0 (not at all reliable) to 10 (completely reliable). Next, participants read the same source information again, but this time together with a claim. For example, having read that ‘Michael is a drug addict’, some participants read the following:

Now imagine that Michael told you the following: ‘One of the best remedies against a severe cough is valium.’

Participants re-rated source reliability on the same Likert scale. No definition of ‘reliability’ was provided. As in Experiment 6.1, each participant saw a script with six texts, with two orders of presentation to control for order effects.

6.4.2 Results & Discussion

This experiment used the same method of analysis as Experiment 6.1 and 2: namely, robust Bayesian parameter estimation

- (1) Expected claims should increase source reliability
- (2) Unexpected claims should decrease source reliability.

The mean change in the perceived reliability of the source as a function of claim expectedness or unexpectedness is shown in Figure 6.7 below. These means suggest are in keeping with (1) and (2): expected claims led to increases in source reliability, unexpected claims to decreases.

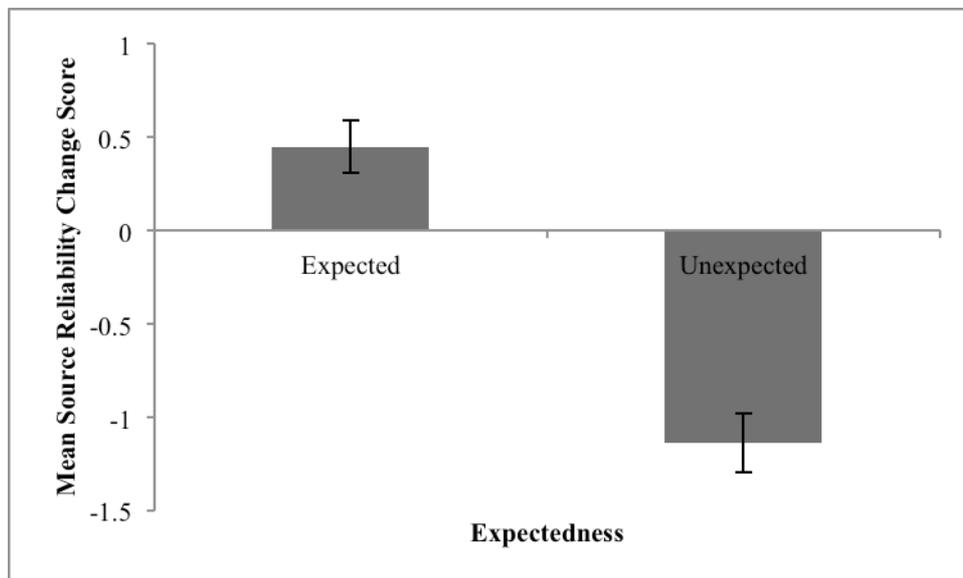


Figure 6.7. Mean change in reliability by expectedness; error bars are standard error

The data were again tested with two one-group analyses with a comparison value of 0.

Expected Claims. The mean estimate for μ was .45 (95% HDI [.18, .74]). The modal estimate for σ was .93 (95% HDI [.75, 1.18]). The modal estimate for $\log_{10}(v)$ was 1.53. The modal estimate for effect size was .49 (95% HDI [.16, 0.79]), which falls outside the conventional ROPE. Figure 6.8 shows the posterior

distribution for effect size and the ROPE. Thus, this analysis shows that expected claims credibly increased source reliability.

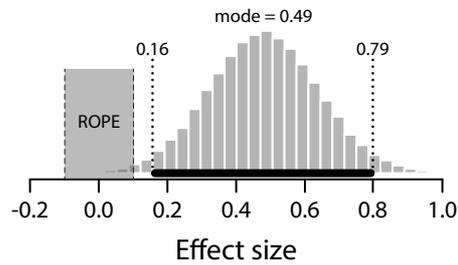


Figure 6.8. Posterior distribution of effect size of reliability change from expected claims. ROPE is -.1 to .1; black bar depicts 95% HDI

Unexpected Claims. The mean estimate for μ was -1.12 (95% HDI [-1.43, -.8]). The modal estimate for σ was 1.37 (95% HDI [1.09, 1.65]). The modal estimate for $\log_{10}(v)$ was 1.16 (95% HDI [.59, 1.96]). The modal estimate for effect size was -.82 (95% HDI [-1.11, -.56]), which falls outside the conventional ROPE. Figure 6.9 shows the posterior distribution for effect size and the ROPE. Thus, this analysis shows that unexpected claims credibly decreased source reliability.

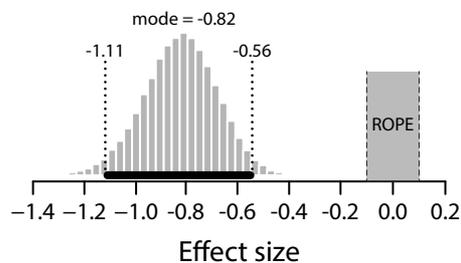


Figure 6.9. Posterior distribution of effect size of reliability change from unexpected claims. ROPE is -.1 to .1; black bar depicts 95% HDI

Summary. These data therefore support predictions (3) and (4). Expected claims increased source reliability; unexpected claims decreased source reliability

As above, the novelty of the data argues for a replication. The same change was made to the materials, namely, dropping the Land Rover item. And, as above,

the replication was run on Amazon Mechanical Turk to allow a more representative sample.

6.5 Experiment 6.4: Replicating reliability change

6.5.1 Methods

Design. The experiments followed the same design as Experiments 6.1 to 6.3: a 2x2 between-subjects design with the following factors: Claim Expectedness (Expected, Unexpected) and Source Reliability (High, Low).

Materials and Procedure. Experiment 6.4 used the same materials and procedure as Experiment 6.3. As for experiment 6.2, two minor changes were made to the materials to adapt them for a predominantly US audience. Firstly, for the Stockholm item, temperatures were given in both Centigrade and Fahrenheit. Secondly, for the nightclub item, US locations were given: Manhattan as an expected location for a prestigious nightclub, and Des Moines, Iowa, as an unexpected location. The Range Rover item was also removed because, in Experiment 6.1, participants' prior beliefs showed that the intended expectedness manipulation had not worked.

Participants. 79 people (31 women; average 35.9) completed online surveys posted on Amazon's Mechanical Turk as a small job (HIT). Participants were recruited and rewarded in the same way as in the experiments above. In addition, a qualification guaranteed that people could not participate if they had previously completed in any of the experiments. As above, one participant's data (not included in the above count) was excluded because that participant reported a first language other than English.

6.5.2 Results & Discussion

The analysis was the same as for Experiment 6.3.

- (1) Expected claims should increase source reliability
- (2) Unexpected claims should decrease source reliability.

The descriptive data, in Figure 6.10, qualitatively match the earlier findings

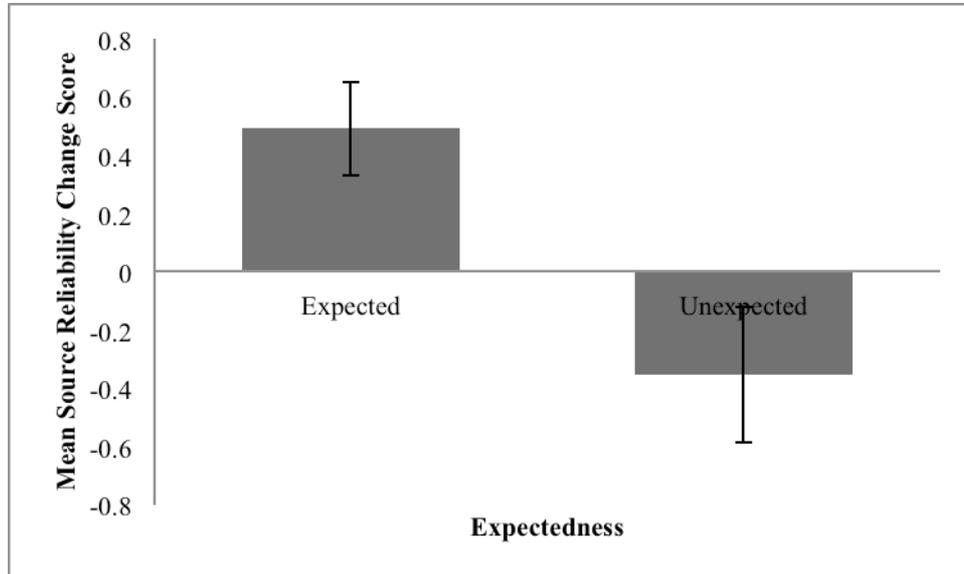


Figure 6.10. Mean reliability change by expectedness; error bars are standard error

Expected Claims. The mean estimate for μ was .51 (95% HDI [.2, .83]). The modal estimate for σ was .89 (95% HDI [.66, 1.19]). The modal estimate for $\log_{10}(v)$ was 1.51. The modal estimate for effect size was .56 (95% HDI [.19, .97]), which falls outside the conventional ROPE. Figure 6.11 shows the posterior distribution of effect size and the ROPE. Thus, this analysis shows that expected claims credibly increased source reliability.

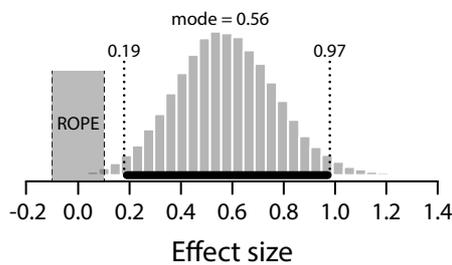


Figure 6.11. Posterior distribution of effect size of reliability change from expected claims. ROPE from -.1 to .1; black bar depicts 95% HDI

Unexpected Claims. The mean estimate for μ was $-.2$ (95% HDI $[-.64, .22]$). The modal estimate for σ was 1.16 (95% HDI $[.65, 1.7]$). The modal estimate for $\log_{10}(v)$ was $.52$ (95% HDI $[.12, 1.77]$). The modal estimate for effect size was $-.18$ (95% HDI $[-.53, .18]$). Although this modal estimate is for a small effect, the 95% HDI includes an effect size of 0. Indeed, the 95% includes the entire ROPE ($-.1$ to $.1$). It is therefore not possible to reject the null hypothesis. But since the 95% also includes effect sizes outside of the ROPE, it is also not possible to confirm the null hypothesis. Figure 6.12 shows the posterior distribution and the ROPE. Accordingly, these data do not allow us to make a decision on the effect of unexpected claims.

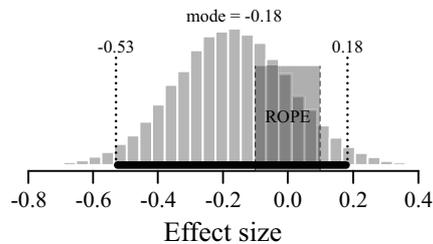


Figure 6.12. Posterior distribution of effect size of reliability change from unexpected claims. ROPE from $-.1$ to $.1$; black bar depicts 95% HDI

Summary. These data support prediction (1) but (in contrast to the preceding experiment) do not allow a decision on prediction (2). Expected claims increased source reliability; but there was no statistical evidence for (or against) unexpected claims decreasing reliability.

The experimental data, thus far, are broadly consistent. They suggest that people use source reliability when assessing claims, and can consider sources to be anti-reliable or negatively correlated with the truth, as predicted by the Olsson and Angere model. The data also suggest that people use message content to judge reliability. The experiments above all tested single claims in a somewhat artificial format, requiring participants to assess a source's reliability twice. Although this pre-

test/post-test structure controls well for individual differences among participants (Slater & Rouner, 1996), such designs are notoriously prone to demand characteristics (Levine & Parkinson, 1994). For the above mechanism to play a major role in belief revision, these beliefs and perceptions should hold across multiple interactions and without the pre-test/post-test structure⁵⁶. The next experiment explores this possibility.

6.6 Experiment 6.5: Story version

Experiment 6.5 adapted the materials from the previous experiments for use in a new paradigm. This paradigm examined the extent to which people spontaneously use message content to revise their beliefs about a source. This paradigm employed a story version of the previous experiments which avoided an explicit source-reliability variable. The idea was to test whether judgments of one claim would impact on judgments of a later claim. Participants read a short text comprising information about a source; a first claim, which varied in expectedness; and a neutral second claim, which participants then rated. If there were systematic differences in the rating of the second claim, this would reflect spontaneous and implicit revision of perceived reliability as the story unfolds. This method approaches a more naturalistic version of the preceding experiments. To accommodate multiple claims, the structure of a trial now read as follows:

⁵⁶ For pre-test/post-test designs, demand characteristics are often explored using a Solomon Four-Group Design (for discussion, see, e.g., Levine & Parkinson, 1994). For present purposes, we could use the following format, illustrated here for the reliability task:

Group 1:	Pre-test	Treatment (Expected Claim)	Post-test
Group 2:	Pre-test	Treatment (Unexpected Claim)	Post-test
Group 3:	No Pre-test	Treatment (Expected Claim)	Post-test
Group 4:	No Pre-test	Treatment (Unexpected Claim)	Post-test

Comparing Groups 1 and 2 would replicate the present data. Comparing Groups 3 and 4 would replicate the data in a post-test only form. Comparing Groups 1 and 3, and 2 and 4, would test whether pre-testing significantly affected the results.

RATIONALITY, PRAGMATICS, AND SOURCES

Imagine you hear Michael, who is a clinical nurse specialist, telling someone “One of the best remedies for a severe cough is valium.” Later, Michael tells you the following: “The new medicine Fluntem can prevent heart attacks and strokes.”

The first claim, to repeat, manipulated expectedness: here, the claim is unexpected.

The second claim was intended to be neutral, with a prior probability of around 0.5.

Participants rated their belief in the second claim on an eleven-point scale of convincingness.

The key prediction was that participants would use the expectedness of the first claim to implicitly judge source reliability, and use this source reliability judgment to inform their judgment of the convincingness of the second claim. Hence, an expected first claim would lead to higher belief in a second claim than would an unexpected first claim. As before, the design was between-subjects: some participants saw expected claims followed by neutral claims (the expected condition), some unexpected followed by neutral claims (the unexpected condition). To aid interpretation of any implicit revision of source reliability in response to expected vs. unexpected claims, we added a third, baseline condition which presented participants only with the neutral, second claim (the null condition).

Given this design, three comparisons are possible: the expected versus null conditions; the unexpected versus null conditions; and the expected versus unexpected conditions. The key prediction is for the expected and unexpected conditions to differ reliably. The null condition is expected to lie between these two, though there are no predictions concerning how far it should be from either, as it reflects how trusting people are initially. If the expected and unexpected conditions

were reliably different from the null condition, this would be a clear replication of the preceding studies.

6.6.1 Methods

Design. There was one independent variable: previous message content (null, unexpected, expected). The dependent variable was claim convincingness on a scale from 0 to 10.

Participants. Two samples were recruited on Amazon Mechanical Turk using the same criteria as for Experiments 6.2 and 6.4; they were again remunerated at a rate of \$.20 per minute. Two separate HITs were posted on different days and at different times of day. Participants were able to participate only in one HIT, and only if they had not previously taken part in Experiment 6.2 or 6.4. The first sample comprised 120 people (45 women; average age 35.98). The second sample comprised 296 people (3 gender non-conformist, 123 female; average age 34.81). Note that, although classical methods do not allow samples to be compiled, Bayesian methods do (Kruschke, 2015).

Materials and Procedure. This experiment used the same set of sources and first claims as Experiments 6.2/6.4. The new, second claim in each pair was as follows: for the valium item, ‘The new medicine Flumentem can prevent heart attacks and strokes’; for the oven item, ‘Pimlico Farm superfine flour is the best on the market for making pasta’; for the horse-racing item, ‘The yacht Azure will beat its competitor Orion at this year’s Cowes Week regatta’; for the Stockholm item, ‘It rained on 13 days in Tübingen, Germany, in May 2013’; for the clubbing item, ‘Kate Siggs is a rising star on the vibrant Australian jazz scene. Participants provided a rating for the second claim on a scale from 0 (not at all convincing) to 10 (completely convincing).

To illustrate the new format:

Imagine you hear Robert, who is a senior sports reporter and has predicted the winner in the last 10 races that he covered, telling someone the following:

“The Australian horse Thunderbolt, who has beaten the British horse Lightening in the majority of the races entered this season, will lose to (beat) Lightening in the upcoming Cheltenham Festival races.”

Later, Robert tells you the following: “The yacht Azure will beat its competitor Orion at this year’s Cowes Week regatta.”

How convincing is this claim about Azure on a scale from 0 (not at all convincing) to 10 (completely convincing)?

6.6.2 Results & Discussion

To analyze the data we averaged the endorsement of the second claim across items to create a mean score for each participant. We then ran the analyses on these scores. There are three relevant analyses: expected condition versus null condition; unexpected condition versus null condition; expected condition versus unexpected condition. As with the one-group analyses, these analyses describe the data with a t-distribution, and estimate the most credible parameter values given the data. For the two-group analyses, the following model applies:

$$P(\mu_1, \mu_2, \sigma_1, \sigma_2, \nu | D) = \frac{P(D | \mu_1, \mu_2, \sigma_1, \sigma_2, \nu) * P(\mu_1, \mu_2, \sigma_1, \sigma_2, \nu)}{P(D)}$$

Subscripts identify group membership. Note that, in this model, there is only one parameter for normality. The technical details are the same as for the one-group analyses. The priors are, likewise, set in the same way. Below, for brevity’s sake, we report estimates for the differences between μ_1 and μ_2 and between σ_1 and σ_2 , for the normality parameter, and for the effect size.

Figure 6.13 shows the descriptive data for the combined data sets.

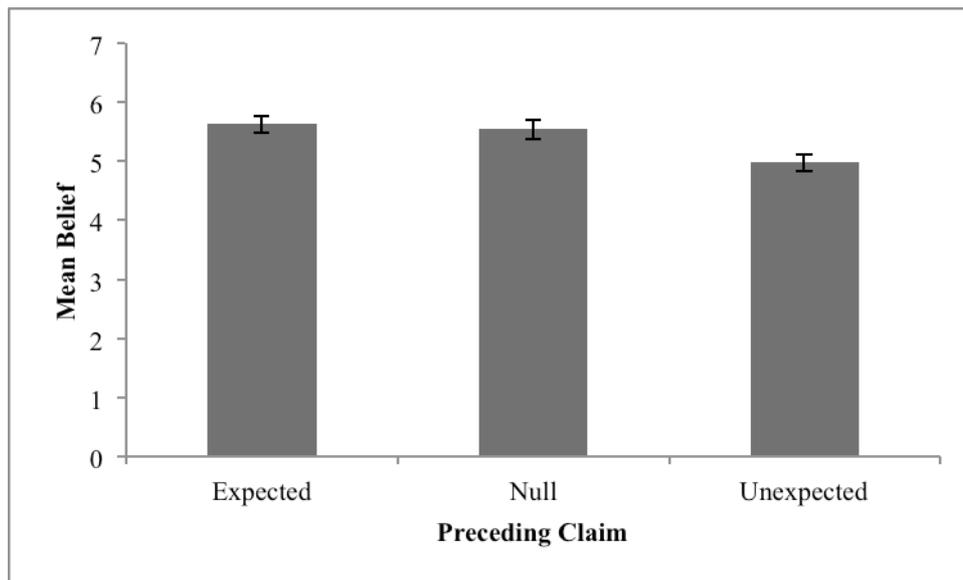


Figure 6.13. Mean belief by expectedness; error bars are standard error

The descriptive data show the predicted pattern in that mean belief was higher in the expected condition, than in the neutral or unexpected condition. The fact that the unexpected condition showed lower mean belief than the null condition is also suggestive of anti-reliability.

Expected v Null. The mean estimate for difference in means ($\mu_{\text{expected}} - \mu_{\text{null}}$) was .09 (95% HDI [-.33, .51]). Note that the 95% HDI includes a difference of zero. The modal estimate for difference in standard deviations ($\sigma_{\text{expected}} - \sigma_{\text{null}}$) was -.25 (95% HDI [-.54, .07]). The modal estimate for v was 1.72 (95% HDI [1.22, 2.17]). The modal estimate for effect size was .06 (95% HDI [-.19, .29]). Figure 6.14 shows the posterior distribution of effect size and the ROPE. Since the 95% HDI for effect size encompasses a conventional ROPE, there is insufficient evidence to determine an effect of preceding expected claims here.

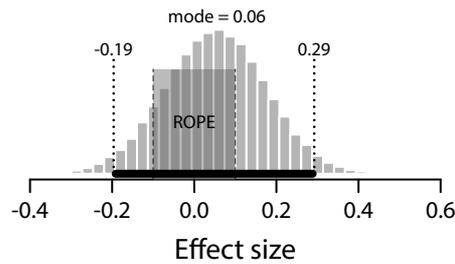


Figure 6.14. Posterior distribution of effect size of difference between expected and null conditions. ROPE from -.1 to .1; black bar depicts 95% HDI

Unexpected v Null. The mean estimate for difference in means ($\mu_{\text{null}} - \mu_{\text{unexpected}}$) was .56 (95% HDI [.12, .98]), suggesting a credible difference in means, with the unexpected condition lower than the null condition. The modal estimate for difference in standard deviations ($\sigma_{\text{null}} - \sigma_{\text{unexpected}}$) was .13 (95% HDI [-.19, .44]). The modal estimate for ν was 1.73 (95% HDI [1.23, 2.18]). The modal estimate for effect size was .32 (95% HDI [.07, .55]). Notice that the 95% HDI excludes an effect size of zero. However, it also overlaps with the conventional ROPE (-.1 to .1) by .3. Figure 6.15 shows the posterior distribution of effect size and the ROPE. Although these data are suggestive, they do not allow us to decide whether or not there is a credible difference between unexpected and null conditions.

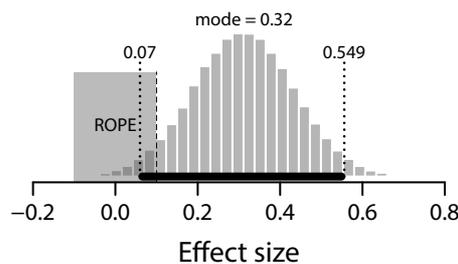


Figure 6.15. Posterior distribution of effect size of difference between unexpected and null claims. ROPE from -.1 to .1; black bar is 95% HDI

Expected v Unexpected. The mean estimate for difference in means ($\mu_{\text{expected}} - \mu_{\text{unexpected}}$) was .64 (95% HDI [.24, 1.05]). The modal estimate for difference in standard deviations ($\sigma_{\text{expected}} - \sigma_{\text{unexpected}}$) was -.1 (95% HDI [-.4, .2]). The modal estimate for ν was 1.63 (95% HDI [1.1, 2.13]). The modal estimate for effect size

was .38 (95% HDI [.14, .63]). The 95% HDI falls outside the conventional ROPE.

Figure 6.16 shows the posterior distribution of the effect size and the ROPE. There is therefore a credible difference between expected and unexpected conditions.

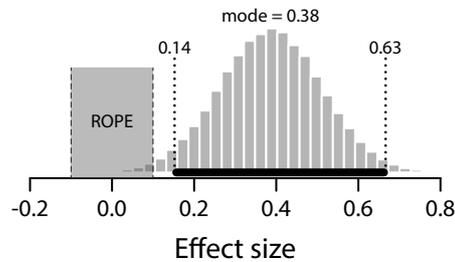


Figure 6.16. Posterior distribution for effect size of difference between expected and unexpected conditions. ROPE from -.1 to .1; black bar depicts 95% HDI

These data, therefore, suggest that people do, indeed, use the expectedness of claims to implicitly judge the reliability of a source, which in turn they use to judge later claims. There is a reliable difference between expected and unexpected claims. There is also a suggestion of anti-reliability, in that the vast majority of the 95% HDI for effect size of the difference between unexpected and null condition fell outside a conventional ROPE. However, since there was overlap, we must avoid a firm decision.

6.7 General Discussion

This chapter set out to explore the relationship between rationality and sources. It did so by taking the case of testimony and exploring how people respond to testimony from partially reliable sources. The aim was to test predictions from two different Bayesian models of testimony. The data offered some support for both models: ratings of convincingness increased with reliable sources; ratings of reliability increased with expected claims, and decreased with unexpected claims. But the data also showed reliable evidence for source anti-reliability – people decreased their belief when there was an unreliable source – which is a behaviour

predicted only by the Olsson and Angere model. Lastly, an experiment suggested a feedback mechanism: the expectedness of a first claim led to an implicit reliability judgement; this judgement then influenced the convincingness of a second claim.

These experiments suggest that people are epistemically vigilant: people monitor the plausibility of incoming claims and the reliability of their sources. The findings are, therefore, broadly compatible with a wide range of research in other domains which offers indirect evidence of vigilance. These domains are the developmental psychology of trust and the psychology of persuasion. The present findings relate even more closely to existing research in Bayesian argumentation. In this nascent field, researchers argue that people are sensitive to the appropriate parameters for Bayesian reasoning about testimony, and that they can be considered at least qualitative, and perhaps noisy quantitative, Bayesian reasoners (Hahn et al., 2009; Harris et al., 2016). What the present findings add is the distinction between two Bayesian models and the clear demonstration that sources can be anti-reliable.

This chapter has evidenced people's sensitivity to a key message characteristic – its expectedness. This finding coheres well with a rich literature in sociolinguistics, as well as data in social psychology, which suggest that people use message characteristics to form impressions of people. For instance, speed gives an impression of competence⁵⁷; choice of register gives an impression of power or powerlessness; lexical diversity gives an impression of competence and control; and moving towards the interlocutor's accent gives an impression of similarity and attractiveness (Holtgraves, 2002). Such characteristics have an impact on persuasion. Slater and Rouner (1996) manipulated messages so that they seemed interesting,

⁵⁷ Though excessive speed makes a speaker seem anxious.

well written, appropriately organized and styled, creative, and so on. Like the present experiments, they had participants assess sources before and after making claims; unlike the present experiments, the claims were extensive persuasive texts. Slater and Rouner found that message quality predicted source assessment, and partially mediated the relationship between initial and ultimate assessments. Message quality also predicted belief change, and partially mediated the relationship between initial source assessment and belief change⁵⁸.

The present experiments offer relevant evidence on the phenomenon of belief (or attitude) polarisation, in which different parties become more extreme in their beliefs as they consider evidence. Belief polarisation is a perennial issue in contemporary political science (Mann & Ornstein, 2013). Participants in the experiments above formed perceptions of reliability on the basis of claims' expectedness, and used these perceptions in evaluating future claims. Participants were also prepared to consider sources anti-reliable, or negatively correlated with the truth. They were prepared to do all of this without knowing about the truth of the claims. These behaviours could be a potent combination. If the behaviours obtain in the real world, then real-world audiences may risk discounting sources solely because they found previous claims unexpected (Hahn & Harris, 2014). This discounting may be still more pronounced if expectedness is determined not just but by our probability estimates but by our likes and dislikes (Osgood & Tannenbaum, 1955). Anti-reliability may provide a valuable complement to existing theories of polarization such as the accounts of backfire effects briefly surveyed in the introduction.

⁵⁸ Slater and Rouner (1996) made other predictions, but these are less relevant for

RATIONALITY, PRAGMATICS, AND SOURCES

From these considerations, a normative question arises: are the behaviours reported above normatively defensible? This chapter has explored two Bayesian models which are supposedly normative. Yet these models make different prescriptions about how to handle information from unreliable sources. Is one model's prescription more rational? An approach to this question is to run simulations with agent-based models. Agent-based models comprise individual agents in an artificial environment (see, e.g., N. Gilbert, 2008). The agents are programmed with a simple set of behaviours, and then interact with other such agents. Agent-based modelling allows the study of complex, non-linear, and emergent behaviours. Of special relevance here, such models can test whether seemingly rational individual behaviour can result in seemingly irrational group behaviours such as polarisation. This approach needs some independent assessment of rationality; such an assessment is not straightforward with polarisation (Jern, Chang, & Kemp, 2014). But it is essential to consider rationality at the group level. In parallel, to assess how rational people are being, we must ask a question that was set aside above: how far do the behaviours generalize? We have seen, for instance, that anti-reliability occurs. But is it the default behaviour or is it triggered in specific situations – for instance, when a source is biased? The normative question should, therefore, be pursued alongside a larger research program exploring testimony across different contexts.

The Bayesian approach to testimony covers similar ground to accounts of persuasion such as the ELM (see the introduction to this chapter). There is, however, a marked difference in style. On the Bayesian approach, a rational agent should

present purposes and were less clearly supported by the data.

RATIONALITY, PRAGMATICS, AND SOURCES

always consider both the claim and the source, and combine this information multiplicatively in updating their beliefs. On the ELM account, an agent considers source information primarily only in the peripheral route. In specific cases, source information may actually form part of the argument; and it may, more generally, influence persuasion through processing effects or metacognition. But sources have, nevertheless, a secondary status in the two identified (i.e. central and peripheral) routes to persuasion. It is hard to test between Bayesian and ELM accounts, in large part because the ELM has so many parameters that it can account for seemingly contradictory data. For instance, the ELM does not, in fact, commit to the existence of only two routes to persuasion; it allows any number of unidentified intermediate routes, compatible with any number of relationships between claims and sources. The ELM does not, however, have as straightforward a fit as the Bayesian models do with the present experiments. The Bayesian models allow – indeed prescribe – a subtle two-way relationship between claims and sources, whether the context has high stakes and high motivation or low stakes and low elaboration. The Bayesian models are also grounded in a deeper, normative account: real-world agents may be behaving as they are because they are trying to process data optimally. There is no room for such normativity in the ELM.

The present chapter has, like Chapters 2 to 4, explored source reliability, but it has produced a different pattern of results. The difference lies in the treatment of information from unreliable sources, and needs explanation. In this chapter, when participants received a claim from an unreliable source, they decreased their belief in the claim. Chapters 2 to 4 presented information in a different way: participants did not provide ratings for a claim before and after assertion. But since some participants rated probabilities without assertion and others with assertion, the data are similar

enough to compare. In Chapters 2 to 4, participants reliably increased belief in response to unreliable sources, which is a different pattern from the current data. However, in Chapters 2 and 3, participants revised their belief only in the conditional probability, which raises the following possibility: people respond differently to source reliability when judging conditional relationships. There are, indeed, hints that this is so. Participants in a reasoning task were more sensitive to source reliability when they were rating a conditional premise than when they were rating a minor premise (Stevenson & Over, 2001). Participants arguably assume that knowledge of conditional relationships is more specialist than knowledge of simple facts: conditional relationships, for example, might map onto subtle features of causal systems (Stevenson & Over, 2001). It is not clear, though, that this point would explain why anti-reliability does not obtain for conditionals.

There is a more straightforward explanation in the way reliability was operationalized. Chapters 2 to 4 manipulated source reliability in a slightly different way from this chapter. In Chapters 2 to 4, unreliable sources were merely comparatively low in the appropriate expertise for a specific scenario. For the most part, the manipulation amounted to simply telling participants who the source was; inexpert sources were intuitively less likely to have the relevant knowledge, but they *could* have it. For instance, in one scenario, an inexpert source was a medical student who claimed that, 'If a patient on this ward has malaria, then they will make a good recovery.' A medical student would surely have some relevant expertise, just less than the professor of medicine in the high expertise condition. In Chapter 6, in contrast, unreliable sources were demonstrably lacking in relevant expertise or accuracy, or were untrustworthy. The materials suggested precisely why the source was deficient for the issue in hand. The Chapter 6 sources were, in other words,

better candidates for anti-reliability. To explore this possibility, Chapters 2 to 4 could be supplemented with new materials. The experiments in this chapter could also be supplemented with more extensive, real-world examples, with more varied manipulations of source reliability. These data would speak to the generalizability of source anti-reliability.

Although this chapter has focused on the link between sources and rationality, the findings also bear on pragmatics. The findings suggest an important distinction between comprehension and acceptance. Participants surely understood the claim that the source was (fictionally) trying to communicate, but they did not blindly accept it. In fact, when they found the source sufficiently unreliable, they decreased their belief in the claim. This distinction is especially striking given that Chapters 2 to 4 stressed the intimate relationship between pragmatics and testimony.

How, then, can we understand the relationship between pragmatics and testimony? A good basis is the definition of communication which underpins contemporary natural-language pragmatics. The exact formalization is disputed (Grice, 1957; Schiffer, 1972; Searle, 1969; Strawson, 1964). But for present purposes the following definition will suffice:

Informative Intention: to inform the audience of something.

Communicative Intention: to inform the audience of one's informative intention.

This informal definition is due to Relevance Theory (B. Clark, 2013, pp. 113-4; see also Sperber et al., 2010; Sperber & Wilson, 1995). The informative intention is the intention to get the audience to believe whatever the speaker is saying; it is fulfilled only if the audience does, indeed, believe what the speaker is saying. The communicative intention is a second-order intention: it refers to the informative

RATIONALITY, PRAGMATICS, AND SOURCES

intention, which is supposed to be overt and to be recognized by the audience. If either intention is fulfilled, then communication is successful. If the informative intention is fulfilled, then the audience believes what the speaker wants them to believe. If only the communicative intention is fulfilled, then the audience understands what the speaker wants them to believe but does not accept it. When both intentions are satisfied, then an audience understands a speaker's claim and accepts it as true in virtue of the speaker's epistemic authority (Sperber et al., 2010). This, in other words, is testimony.

This framework offers a subtle way of explaining the intimate connection between testimony and pragmatics. Testimony requires pragmatics but emphasizes an assessment of epistemic authority (source reliability) and believability (expectedness) (Sperber et al., 2010). The word 'emphasizes', here, is key, for assessments of reliability and believability may also play a role in pragmatics; this role is simply less pronounced. Reliability is arguably required for pragmatics in the following sense: hearers need an initial stance of trust, a preparedness to change their beliefs, to entertain the cooperation required for demanding pragmatic inferences; and if this trust drops below a certain level, then pragmatics may simply stop (McCready, 2014; Sperber et al., 2010). Believability is, according to Sperber et al., also important in pragmatics. Indeed, Sperber et al. argue that the same mechanisms underpin the assessment of believability and interpretation. Both require the assessment of new information relative to background knowledge. Interpretation, they argue, occurs through a search for cognitive effects: new conclusions in light of new information and contextually activated beliefs; increased or decreased confidence in such beliefs; or the contradiction and revision of contextually activated

RATIONALITY, PRAGMATICS, AND SOURCES

beliefs. Interpretation identifies, in other words, inconsistencies with current beliefs; testimony is the business of dealing with these inconsistencies.

The foregoing discussion gives rise to certain predictions. Firstly, pragmatic inferences are not automatic; they require trust. Thus, systematic unreliability should decrease the number of pragmatic inferences. Secondly, if, as theory and the present data suggest, there is epistemic vigilance towards sources and content, then speakers may have tools available to them to manage their reputations, so that they can maintain cooperation, hence pragmatics, while making claims for which they lack complete confidence (McCready, 2014; Sperber et al., 2010). The next chapter will address one proposed mechanism.

7 Evidential Language

This chapter returns to the relationship between pragmatics and sources. Chapter 6 presented novel data which argue for vigilance to both the source and content of a message. This vigilance seems intuitively rational, but raises the possibility that full pragmatic communication simply stops when a source has given unexpected information, because reliability and attendant trust drop below some threshold. In other words, pragmatic inference depends on source factors. This chapter tests a mechanism which has been proposed to allow communication to continue in such contexts: hedging (McCready, 2014).

The experiments in Chapter 6 form the basis of this chapter. Chapter 6 suggested that people take the perceived reliability of sources into account when revising their beliefs in response to testimonial evidence for a claim. The experiments also suggested that people spontaneously use the degree of (mis)match between a claim and their present priors for the claim to revise their beliefs about the source. This revision occurred even when no overt judgments of reliability were required, suggesting a fundamental bi-directional relationship between content and source in the processing of testimonial evidence. Such behaviour fits, at least qualitatively, with Bayesian accounts of testimony (e.g., Olsson & Vallinder, 2013) within the formal epistemology literature.

When speakers make claims, their reliability is under threat, with damaging consequences. On the one hand, speakers are expected to give informative, relevant information (see, e.g., Grice, 1975; McCready, 2014). On the other, if speakers give information which turns out to be false – or even merely seems implausible - they face reputational damage which may impact on future cooperation and undermine

the very basis of communication (McCready, 2014). How then do people maintain trust and cooperation given that none of us, realistically, is perfectly reliable?

McCready (2014) argues that reputation management is crucial to communication. He argues, further, that a crucial strategy is hedging, that is, qualifying assertions in such a way that they allow for exceptions. Take the following examples, where the hedges are italicized:

(a) John is *sort of* stupid

(b) *I suspect that* it is cold outside

(c) *I might be wrong, but* Palin is not going to be elected.

(d) *This might not be true, but* she doesn't really care about you.

(McCready, 2014, p. 39)

These hedges all soften the assertions. But while (a) and (b) seem to modify the assertions rather directly, (c) and (d) seem to work more indirectly: a bold assertion is made, but a 'shield hedge', or disclaimer, is added (McCready, 2014).

A wide range of evidential language could, in principle, function as hedges. Some languages, for instance, possess closed evidentiality systems. These are grammatical categories with restricted membership which convey degrees of evidence such as direct evidence from perception, indirect evidence inferred from perception, and evidence from hearsay (Matsui & Fitneva, 2009; Simons, 2007). Languages such as English possess open systems, which convey evidentiality in less routine ways through the lexicon (Matsui & Fitneva, 2009). For example, English speakers can signal evidentiality through verbs such as 'see', 'hear' and 'feel' (Simons, 2007). Similarly useful for hedging is the expression of speaker certainty, for instance, through modal expressions such as 'possibly', 'probably' and 'certain'

and propositional attitudes like ‘suspect’, ‘think’ and ‘am/is/are sure’ (Portner, 2009; Simons, 2007).

The experiments reported below investigate whether hedging protects a speaker’s reputation, as suggested by McCready (2014). The experiments were simple belief-change tasks based on the source reliability task from Collins, Hahn, von Gerber, and Olsson (2015, 2017) reported in the previous chapter. Participants read an initial description of a source, whose reliability they then rated on a scale from 0 (not at all reliable) to 10 (completely reliable). They then read the source making a claim, after which they rated the source’s reliability again on the same scale. In a between-subjects design, some participants read unmodified claims; some read claims modified by a hedge; in Experiments 7.1 to 7.3 a final group read claims modified by ‘am certain’. Across the groups, we fixed the sources as reliable, to avoid floor effects, and fixed the claims as unexpected, since it is with unexpected claims that a downward revision of source reliability occurs.

Experiments 7.1 to 7.3 included the strong claims, modified with ‘am certain’, on the following grounds. There are a wide range of propositional attitudes, including weaker ones, such as ‘suspect’, ‘suppose’, ‘think’ and ‘believe’, and stronger ones, such as ‘am certain’, ‘am sure’, ‘am convinced’ and ‘know’. It is more informative to consider the full range of propositional attitude strength. If people use propositional attitudes to calibrate source reliability, then, for consistency, weaker propositional attitudes should protect speakers against downward revisions and stronger propositional attitudes should increase the size of the downward revisions.

This chapter reports the results of four experiments. These experiments all test the hypothesis that hedging with propositional attitudes protects speakers against the downward revision of source reliability. More specifically, they test the

hypothesis that, relative to a null condition, hedging should attenuate the downward revision of reliability. Where the strong ('am certain') condition is included, the hypothesis is that this propositional attitude will worsen the downward revision of reliability.

7.1 Experiment 7.1

7.1.1 Methods

Participants. 160 participants completed web surveys; the data were retained for the 159 native English speakers (84 female; average age 37.35) who completed the web surveys. This sample size was selected to give high power (.8) for a medium effect size.

Materials. There were five items which took the following form:

Initial source information: Michael is a clinical nurse specialist.

Claim: Now imagine that Michael told you the following: 'One of the best remedies for a severe cough is valium.'

The item above corresponds to the null condition. In the 'strong claim' condition, participants read the claim embedded under 'I am certain that...' In the 'weak claim' condition, participants read the claim embedded under the 'I suspect that'.

Since the materials in Chapter 6 reliably produced an effect, we adopted items from those studies: namely, the reliable sources and the unexpected claims. A clinical nurse specialist claimed that valium was one of the best remedies for a severe cough. A baker claimed that the varying temperatures of a particular oven made it perfect for fluffy and crispy bread. A journalist with a good track record predicted that a horse would beat a competitor despite losing to that horse in the majority of races. A retired meteorologist claimed that the maximum summer

temperature in Stockholm in June 2013 was 15 degrees. And a respected DJ claimed that a club in Detroit had the reputation of one of coolest in the world.

The propositional attitudes ‘am certain’ and ‘suspect’ were chosen after a pre-test of 36 propositional attitudes for a separate study, not reported in this thesis. In this pre-test, the words were randomly assigned to three separate lists to be rated by three corresponding groups of participants (20 per group). Participants read a neutral claim, such as ‘Louise has left town’, embedded under a propositional attitude. Their task was to compare the strength of the claim to a fixed reference point, namely the claim modified by the non-factive verb ‘think’. The response scale range from 0 to 10, where 0 was labelled ‘much weaker’, 5 as ‘the same’ and 10 as ‘much stronger’. The propositional attitudes ‘am certain’ and ‘suspect’ were chosen because they were far apart ($M = 9.48$ and $M = 5.18$, respectively); moreover, ‘suspect’ specifically features in McCready’s (2014) analysis.

Procedure. The experiments were posted on Amazon Mechanical Turk via the intermediary MTurk Data (www.mturkdata.com). We set high qualifications for the task: participants had to be resident in the US and have an overall approval rating of 99% for 1,000 previously completed tasks. These criteria effectively limited participation to native English speakers. On the consent page, participants were told that the task would assess how people judge information they receive from other people. No information was given about the specific manipulation. After giving informed consent, participants were assigned, round-robin, to a condition. They were shown the following instructions:

Thank you for taking part in this study. You’ll be shown some descriptions of people, and will be asked to indicate how reliable these people seem, by

selecting the appropriate number on a scale from 0 (not at all reliable) to 10 (completely reliable).

Participants then saw 5 items, comprising initial source information and then a claim. Participants gave a prior rating after the initial source information, and a posterior rating after the claim. The order of presentation was counterbalanced. Finally, participants were given debriefing information, and were paid \$0.75, a fee chosen to exceed the US minimum wage for reasonable completion times.

7.1.2 Results

First, change scores were calculated by subtracting the prior rating from the posterior rating. Change scores were then averaged across items to produce a single score per participant⁵⁹.

Figure 7.1 shows the descriptive data. These data do not show the expected pattern of data. In all conditions, there is downward revision of source reliability and, indeed, the ‘suspect’ condition prompts the biggest downward revision of source reliability.

⁵⁹ Note that this design and analysis were chosen for consistency with the previous chapter. The same restrictions apply: we must be cautious about generalizing beyond the existing items. However, McCready’s (2014) theory should apply to the present items. His theory at least requires a major restriction to explain why the present items should be exempt. Chapter 6 suggested a more generalized design for exploring source anti-reliability across contexts. This design could equally be used to test McCready’s theory.

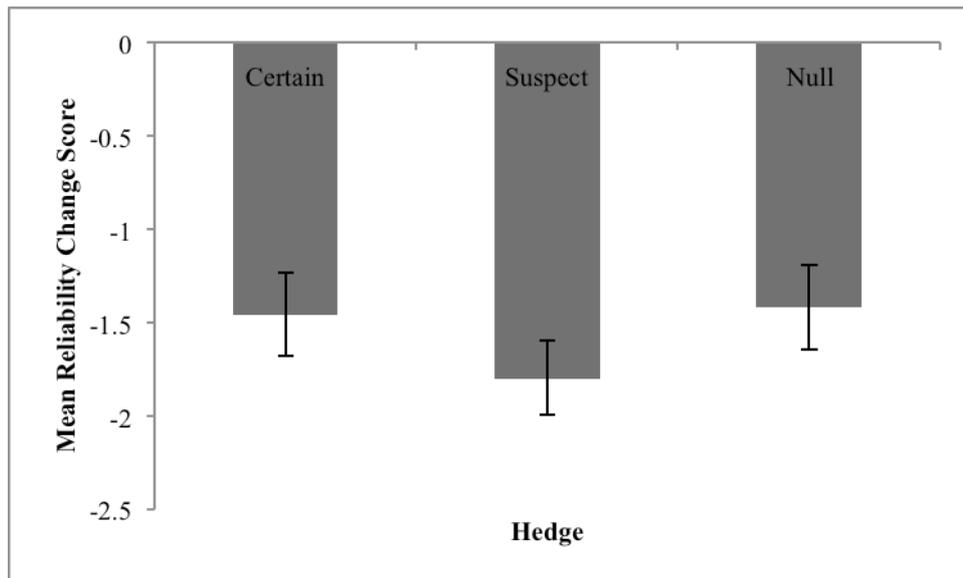


Figure 7.1. Mean change scores by condition; error bars are standard error

Inferential statistics support this picture. To correct for multiple comparisons, the significance level was set at $p = .006$ (the same applies to Experiments 7.2 and 7.3). A one-way ANOVA showed no significant effect of propositional attitude, $F(2,156) = .91, p = .40, \text{partial } \eta^2 = .01$. One-sample t-tests confirmed that, in each condition, there was a significant decrease in ratings of source reliability: in the ‘certain’ condition, $M = -1.46, 95\% \text{ CI } [-1.92, -.99], t(50) = -6.24, p < .001, d = .87$; in the ‘suspect’ condition, $M = -1.80, 95\% \text{ CI } [-2.20, -1.40], t(53) = -8.99, p < .001, d = 1.22$; and in the null condition, $M = -1.42, 95\% \text{ CI } [-1.88, -.96], t(53) = -6.23, p < .001, d = .85$.

Since there were outliers in each condition, we followed the guidelines of Wilcox (2016) and checked the above findings against robust versions of the tests, using the WRS2 package (Mair, Schoenbrodt, & Wilcox, 2017). These analyses used 20% trimmed means. A robust one-way ANOVA (‘t1way’ function) yielded a non-

significant result of propositional attitude, $F(2, 63.43) = 1.88, p = .16, \zeta = .19$.⁶⁰

Robust one-sample analyses were carried out using the function ‘trimpb’, which calculates confidence intervals for trimmed means using a percentile bootstrap method. In each condition, these analyses indicated significant decreases in ratings of source reliability: in the ‘certain’ condition, $M_{trimmed} = -1.32, 95\% \text{ CI } [-1.75, -0.90], p < .001$; in the ‘suspect’ condition, $M_{trimmed} = -1.82, 95\% \text{ CI } [-2.19, -1.41], p < .001$; and in the null condition, $M_{trimmed} = -1.34, 95\% \text{ CI } [-1.82, -0.92], p < .001$.

7.1.3 Discussion

The data above do not support the idea that a key function of hedges is the protection reputation (McCready, 2014) as no protective effects were obtained. There is no evidence of a reliable difference between conditions, and all claims, whether modified by a propositional attitude or not, prompted a significant downward revision of source reliability. It is precisely such a downgrade in reliability that McCready (2014) suggests hedging protects against.

7.2 Experiment 7.2

This experiment attempted to replicate the findings of Experiment 2 while ruling out the possibility that participants did not process the hedges themselves to sufficient depth. Two changes were made to the design. Firstly, we used a simple attentional device, putting the propositional attitudes in capitals. Secondly, we informed participants that they may have to perform a memory test on the items.

⁶⁰ ζ is a robust explanatory measure of effect size, and can be understood in the same way as r (Mair, Schoenbrodt, & Wilcox, 2017). Note that the effect size given, here, is not trivial, but the pattern is not as McCready (2014) would predict, with the largest antireliability being for ‘suspect’.

Participants in the ‘certain’ and ‘suspect’ conditions were, indeed, tested for whether they could remember the propositional attitudes used after rating all items.

7.2.1 Methods

Participants. 170 participants completed web surveys, the same selection criteria having been used. The data were retained for 165 participants (96 female; average age 39.93) who passed the manipulation check (see Procedure). As above, this sample size gave us high power (.8) to detect a medium-sized effect.

Materials. The materials were identical to those in Experiment 1, except that the propositional attitudes were in capitals.

Procedure. The procedure was also identical to that of Experiment 1 except for the addition of the memory test. Participants were informed about the memory test in the instructions, which were otherwise as in Experiment 1. The memory test occurred at the end of the experiment, and in a fixed order. Firstly, participants performed a recall test, which had the following instructions:

On the previous pages you’ve seen people making claims. For example, one claim was ‘I.....that one of the best remedies for a severe cough is valium’, where the ‘...’ stands for a missing word or words. These claims were always introduced by the same word or words (in capital letters).

They were then asked to type the word into a text box, or to type ‘don’t know’ if they could not remember. Secondly, participants performed a recognition test. The above instructions were repeated (introduced by ‘As you’ve just read’), and this time participants had to choose the correct word from a drop-down list. This list included the fillers ‘doubt’ and ‘am uncertain’. The order of this list was counterbalanced across participants. Finally, participants answered the following sincerity question, based on Aust, Diederhofen, Ullrich, and Musch (2013):

‘And, finally, it would be very helpful if you could tell us whether you have taken part seriously, or whether you were just clicking through to take a look at the survey. Please note, you will still get paid however you answer. Did you take part seriously?’

Participants selected their response from a drop-down list with the options ‘Yes – Keep my data’ and ‘No – Please throw my data away’.

7.2.2 Results

The recognition test was chosen as the minimum threshold for keeping the data. Five participants data were rejected for failing to pass the test. The threshold was waived in two cases, where participants in the ‘certain’ condition correctly recalled ‘am certain’ but then chose ‘suspect’ from the drop-down list. Since recall is the harder test, keeping the data is justified; these participants may have taken the repeated question to mean that they had made an error, and then selected the closest item from the drop-down list.

The recall test provides a more stringent threshold. Approximately 76% of participants passed the recall test, rising to 84% if we include near synonyms, such as ‘know’, ‘am sure’ and ‘guarantee’ for ‘am certain’ and ‘hear’, ‘believe’ and ‘think’ for ‘suspect. However, removing those participants who only passed the recognition test made no difference to the significance of the results. We therefore report the analyses on all participants who passed the recognition test.

As Figure 7.2 shows, the descriptive data suggest a similar pattern to Experiment 1, in that all conditions show substantial negative change scores, of comparable magnitude to above. Note, though, that ‘suspect’, this time, receives the smallest change scores.

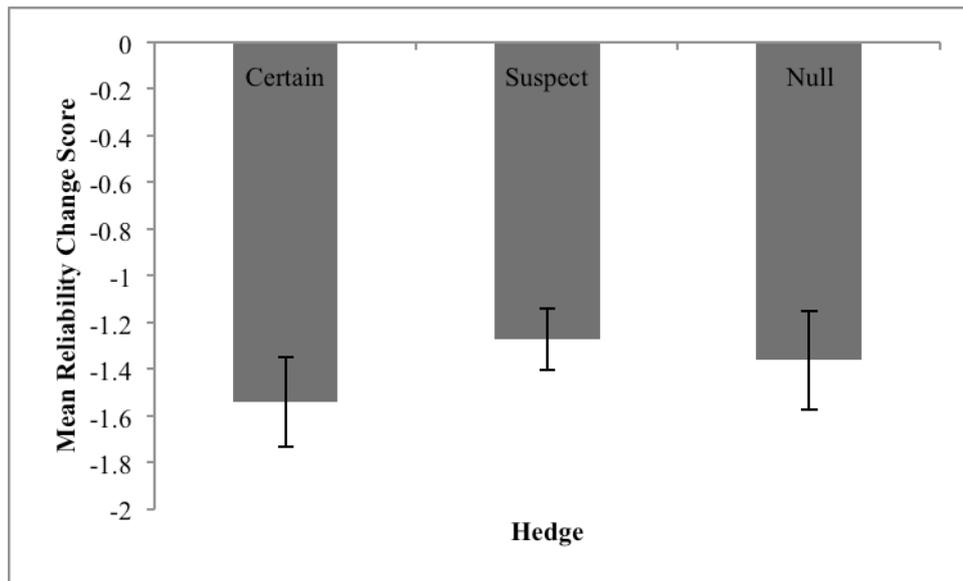


Figure 7.2. Mean change scores by condition; error bars are standard error

A one-way ANOVA supports this picture, and replicates the analyses above. There was no significant effect of propositional attitude, $F(2,162) = .51, p = .60$, partial $\eta^2 = .01$. One-sample t-tests confirmed that, in each condition, there was a significant decrease in ratings of source reliability: in the ‘certain’ condition, $M = -1.54$, 95% CI [-1.95, -1.14], $t(53) = -8.06, p < .001, d = 1.10$; in the ‘suspect’ condition, $M = -1.28$, 95% CI [-1.55, -1.01], $t(51) = -9.66, p < .001, d = 1.34$; in the null condition, $M = -1.36$, 95% CI [-1.76, -.96], $t(58) = -6.53, p < .001, d = .85$.

Since there was an outlier in the null condition, and the suggestion of non-normality in the residuals, the findings were checked against robust analyses using 20% means. A robust one-way ANOVA (‘bwtrim’ function) produced a non-significant effect of propositional attitude, $F(2,66) = .38, p = .68, \zeta = .09$. As above, robust one-sample analyses were performed by calculating the 95% confidence intervals for the trimmed means. In each condition, these analyses indicated significant decreases in ratings of source reliability: in the ‘certain’ condition, $M_{trimmed} = -1.47$, 95% CI [-1.92, -1.02], $p < .001$; in the ‘suspect’ condition, $M_{trimmed} =$

-1.23, 95% CI [-1.54, -.91], $p < .001$; and in the null condition, $M_{trimmed} = -1.25$, 95% CI [-1.69, -.84], $p < .001$.

7.2.3 Discussion

Experiment 2 replicated Experiment 1 even though it drew attention to the propositional attitudes, and even for participants who passed the memory test. There was, once again, no significant effect of propositional attitude – indeed, the effect sizes were smaller in Experiment 2 – and, contrary to the idea of shielding, ratings of source reliability dropped significantly in all conditions. These two data sets suggest, then, that hedging with propositional attitudes does not protect a speaker’s reputation when making unexpected claims.

7.3 Experiment 7.3

The data above are suggestive. But they rely on an assumption: that participants judged the strength of the propositional attitudes in the same way as the participants in the pre-test. It is possible (but perhaps unlikely) that, because of variation in dialects, the participants in Experiments 7.1 and 7.2 simply did not perceive enough of a difference between ‘am certain’ and ‘suspect’ to refrain from downgrading the sources’ reliability. This experiment extends the design of Experiment 7.2 by adding a manipulation check in which participants rated the strength of claims of the form ‘X happened’, ‘I suspect that X happened’, and ‘I am certain that X happened’.

7.3.1 Methods

Participants. 161 participants completed web surveys, this number being chosen to give high power (.8) to detect a medium-sized effect. The data were retained for 151 participants (75 female; average age 37.69) who were native speakers of English and passed the manipulation check (see Procedure).

Materials. The materials were identical to those in Experiment 7.2.

Procedure. The procedure was also identical to that of Experiment 7.2 except for the addition of a manipulation test after the memory test. For the manipulation test, participants were given the following instructions:

Below there are three statements. ‘X’, in the statements, stands for any event.

Please rate how strong these statements are on a scale from 0 to 10, where 0 is ‘very weak’ and 10 is ‘very strong’.

Participants then rated the sentences ‘X happened’, ‘I suspect that X happened’, and ‘I am certain that X happened’; the order of these sentences was counterbalanced across surveys.

7.3.2 Results

First, change scores were calculated by subtracting the prior rating from the posterior rating. Change scores were then averaged across items to produce a single score per participant. Figure 7.3 shows the descriptive data. These data replicate the patterns in Experiments 7.1 and 7.2. In all conditions, there is downward revision of source reliability.

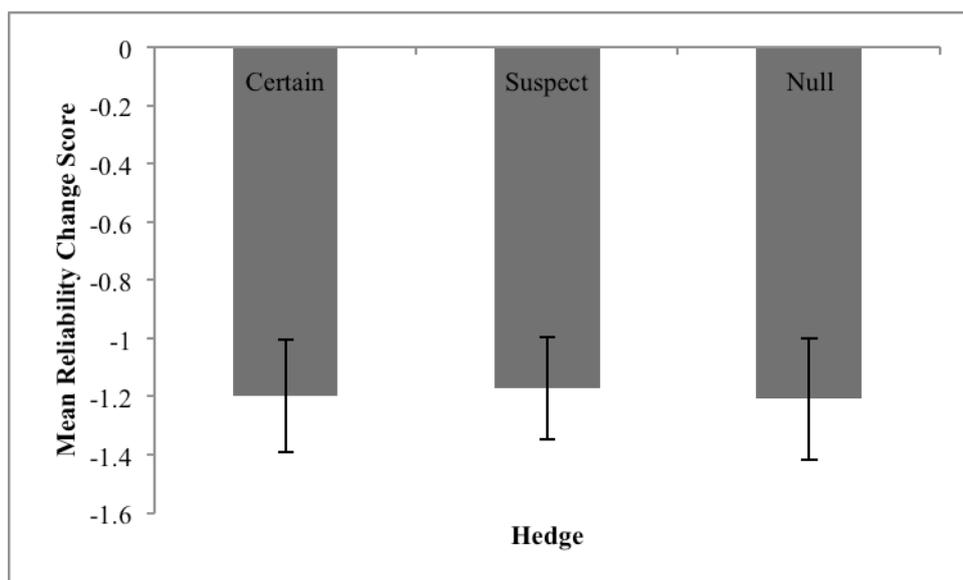


Figure 7.3. Mean change score by condition; error bars are standard error

RATIONALITY, PRAGMATICS, AND SOURCES

Statistical analysis supported this interpretation. A one-way ANOVA showed a non-significant effect of condition, $F(2,148) = .03, p = .97$, partial $\eta^2 = .0004$. One-sample t-tests showed that there was a significant decrease in perceived reliability for the ‘certain’ condition ($M = -1.24$, 95% CI [-1.63, -.85], $t(51) = -6.40, p < .001, d = .89$), ‘suspect’ condition ($M = -1.17$, 95% CI [-1.53, -.79], $t(46) = -6.64, p < .001, d = .97$), and the null condition ($M = -1.21$, 95% CI [-1.62, -.79], $t(51) = -5.83, p < .001, d = .81$). Since there were outliers in each condition, the conventional analyses were replicated using robust equivalents based on trimmed means. A robust one-way ANOVA showed a non-significant effect of condition, $F(2, 58.88) = .04, p = .96, \zeta^2 = .04$. As above, robust one-sample analyses were performed by calculating the 95% confidence intervals for the trimmed means. In each condition, these analyses indicated significant decreases in ratings of source reliability: in the ‘certain’ condition, $M_{trimmed} = -1.18$, 95% CI [-1.53, -.86], $p < .001$; in the ‘suspect’ condition, $M_{trimmed} = -1.18$, 95% CI [-1.47, -.87], $p < .001$; and in the null condition, $M_{trimmed} = -1.11$, 95% CI [-1.55, -.74], $p < .001$.

Figure 7.4 shows the descriptive data for the manipulation check, which suggested that ‘suspect’ was, indeed, rated as considerably weaker than ‘am certain’ and an unmodified assertion. There was little difference between ‘am certain’ and the unmodified assertion.

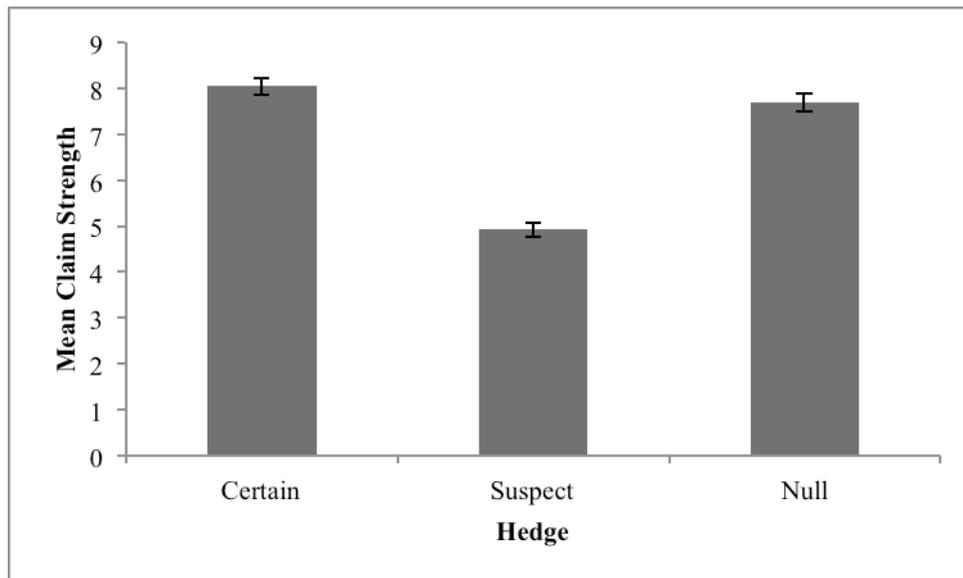


Figure 7.4. Mean claim strength; error bars are standard error

A one-way within-subjects ANOVA (with Greenhouse-Geisser correction) supported this picture. There was a significant effect of word, $F(1.76, 264.27) = 115.5, p < .001$, partial $\eta^2 = .44$. Pairwise comparisons were performed with the Bonferoni adjustment for multiple comparisons. ‘Suspect’ was rated on average 3.13 (95% CI[2.53, 3.74]) lower than ‘certain’, which difference was significant ($p < .001$). ‘Suspect’ was rated on average 2.78 (95% CI [2.20, 3.36]) lower than the unmodified assertion, which difference was also significant ($p < .001$). ‘Am certain’ was rated on average .35 (95% CI[-.79, .09]) higher than the unmodified assertion; this difference was not significant ($p = .16$). These analyses did not require confirmation with robust methods.

7.3.3 Discussion

Experiment 7.3 replicated the effects of Experiments 7.1 and 7.2 while demonstrating that participants were, indeed, sensitive to the difference in strength of the hedges. These data are further evidence that hedging with propositional attitudes does not protect a source’s reputation when making unexpected claims.

7.4 Experiment 7.4: Shield Hedges

The preceding three experiments raise doubts about speakers' ability to protect their reputation: even hedged claims produced a decrease in perceived source reliability, a finding which is problematic for McCready's (2014) account.

Nevertheless, McCready distinguishes between hedging with propositional attitudes and shield hedging, that is, hedging with phrases such as 'I might be wrong, but...'

On his account, propositional attitudes directly weaken the propositional content, whereas shield hedges do not; they merely frame it. It is worth testing, then, whether shield hedges are more effective at protecting reputation, operationally defined as source reliability.

7.4.1 Methods

Participants. 128 participants completed web surveys, the same selection criteria having been used. The data were retained for 124 participants (71 female, 1 non-binary; average age 38.36) who passed the manipulation check and a sincerity question (see Procedure). As above, this sample size gave us high power (.8) to detect a medium-sized effect.

Materials. The materials used the same claims as in the previous experiments. In this experiment, there were two conditions: a shield-hedged condition, in which claims were preceded by 'I might be wrong, but...'; and a null condition, in which claims were presented as unhedged assertions.

Procedure. The procedure was also identical to that of Experiment 7.3 except for a modification to the memory test. Since the hedge phrase was more complex, only the recognition task was used. All participants passed this test.

7.4.2 Results

Figure 7.5 reports the descriptive data by condition. There is no evidence that hedging protected sources' reputations. Indeed, reliability ratings are lower in the hedged than in the unhedged condition.

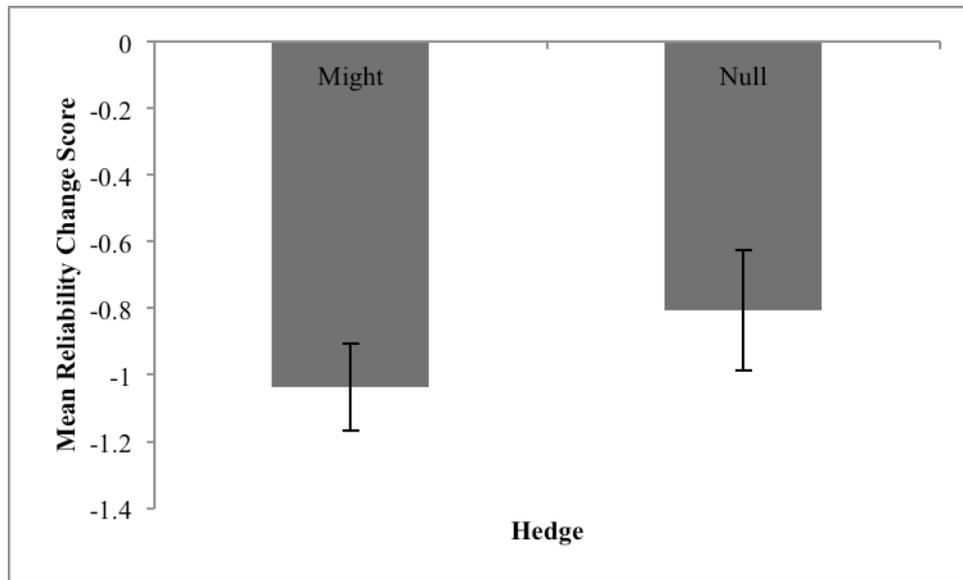


Figure 7.5. Mean change scores by condition; error bars are standard error

To correct for multiple comparisons, inferential tests used the significance level $p = .008$. A between-subjects t-test (equal variances assumed) showed a non-significant difference between the conditions, $M = -.23$, 95% CI $[-.67, .21]$, $t(123) = -1.04$, $p = .30$, $d = .18$. One-sample t-tests confirmed that the hedged condition ($M = -1.04$, 95% CI $[-1.30, -.77]$, $t(66) = -7.89$, $p < .001$, $d = .96$) and unhedged condition ($M = -.81$, 95% CI $[-1.17, -.44]$, $t(57) = -4.43$, $p < .001$, $d = .58$) both showed significant decreases in perceived reliability. The presence of outliers justified the use of supporting robust analyses. A robust between-subjects t-test showed a non-significant difference in 20% trimmed means, $t(65,96) = 1.15$, $p = .26$, $\zeta = .15$. Robust equivalents of one-sample t-tests were performed by estimating the 95% CI intervals on the trimmed means using the percentile bootstrap method. The hedged condition showed a significant decrease in source reliability, $M_{trimmed} = -1.01$, 95%

CI [-1.31, -.72], $p < .001$, as did the unhedged condition, $M_{trimmed} = -.73$, 95% CI [-1.12, -.36], $p < .001$.

Figure 7.6 shows the descriptive data for the manipulation check. The ratings for ‘I might be wrong, but’ are comparable to those of ‘suspect’ and substantially lower than the unhedged control.

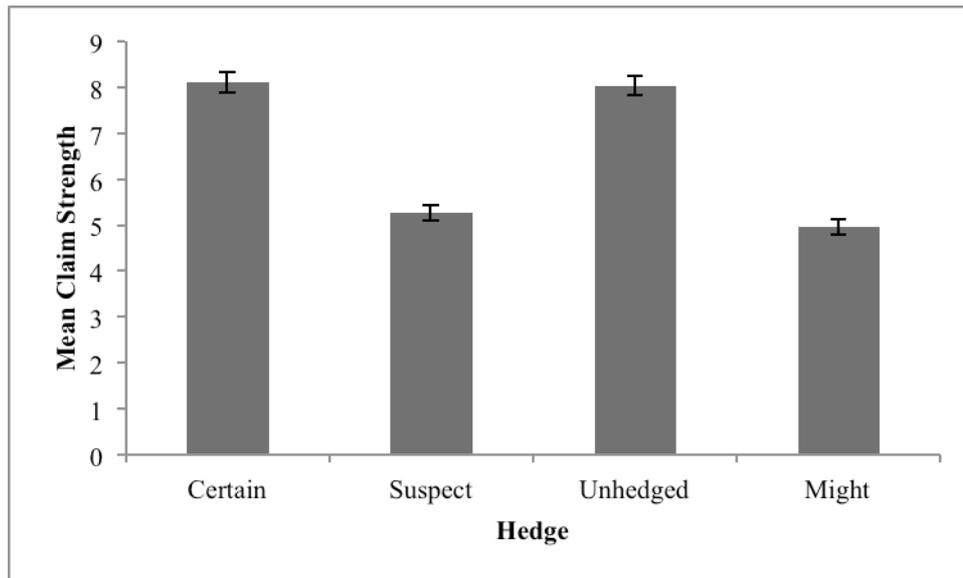


Figure 7.6. Mean claim strength (manipulation check); error bars are standard error

A one-way repeated measures ANOVA (with Greenhouse-Geisser correction) showed a significant effect of word, $F(2.18, 270.03) = 109.33$, $p < .001$, partial $\eta^2 = .47$. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. ‘Certain’ was rated on average 2.83 higher than ‘suspect’ ($p < .001$); only .08 higher than the unhedged claim ($p = 1$); and 3.14 higher than ‘I might have been wrong, but...’ ($p < .001$). The unhedged claim was rated on average 2.75 higher than ‘suspect’ ($p < .001$) and 3.06 higher than ‘I might have been wrong, but’ ($p < .001$). ‘Suspect’ was rated on average .31 higher than ‘I might be wrong’ ($p = .55$). These analyses did not require confirmation with robust methods.

7.4.3 Discussion

The data above offered no evidence for a protective effect of shield hedging; in both conditions, there was a significant decrease in perceived reliability. This finding adds further evidence against McCready's (2014) account and, given the theoretical centrality of shield hedging, is still more problematic.

7.5 General Discussion

We receive much of our information about the world through the testimony of other people. We do not seem to accept such information uncritically; rather, we are vigilant to the reliability of the source and the content of the messages we receive. There is a tension between this vigilance and the trust putatively required for cooperative communication (McCready, 2014). How can we make assertions, lacking full justification, without fatally undermining our reputations and, consequently, trust? This chapter has explored one plausible way in which this tension can be resolved: hedging, with propositional attitudes or shield hedges.

McCready (2014) suggests that speakers are careful to manage their reputations, especially in repeated interactions, and that hedging provides an important strategy for so doing. Although this account is intuitively appealing, and forms part of an impressive theory of communication, the present data suggest that hedging is not enough for reputation management. Possibly, other types of hedging might be more effective, but it is hard to imagine a design which would make hedging much more prominent than above. Future work should nevertheless examine this possibility. Such work should also examine whether hedging might work in contexts where the choice of hedge is clearly contrastive. In Experiment 7.2, participants were specifically alerted to the propositional attitudes, but had nothing

RATIONALITY, PRAGMATICS, AND SOURCES

to compare the particular propositional attitude to. Hedging may, in fact, rely on pragmatic inference. Take, for instance, the following:

Bob: Looks like it'll rain tomorrow.

Sandy: Are you sure?

Bob: Well, I suspect it will.

In this context, Bob may be taken to be implicating that he is not sure: that he only has sufficient confidence to use 'suspect' (Horn, 1989; Levinson, 2000; Van Der Auwera, 1996; Verstraete, 2005). But if hedging works only in a context where there is a clear pragmatic component, it presumably plays less of a role in justifying and maintaining cooperation, since pragmatics is typically taken to be the result of cooperation, not a precondition for it.

There are at least two ways to explore whether hedging relies on pragmatics. One is to develop materials such as the example dialogue above, where a demand for certainty is met with a hedged claim. Another is to switch to oral presentation. It seems intuitively likely that speakers can hedge more successfully by using contrastive stress on the hedge, such as 'I SUSPECT that Valium is one of the best remedies for a severe cough'. Contrastive stress draws attention to a set of alternatives - here, alternatives to the verb 'suspect' - and may thereby prompt a modal scalar implicature to 'not certain'. Since it is natural to read capitalised text with contrastive stress, participants may have generated their own contrastive stress in Experiments 7.2 and 7.3, which used capitalisation as an attentional cue. Contrastive stress is not, however, obligatory with visual presentation; only oral presentation would make this stress pattern unambiguous.

Other aspects of the design might also have contributed to the lack of an effect: in particular, how the design operationalized unexpectedness and reliability.

RATIONALITY, PRAGMATICS, AND SOURCES

The unexpected claims could simply have been too unexpected to be blocked; with more moderate claims, hedging could have been more successful. While this point argues for replication with new materials, the unexpected claims were not wildly implausible. They were implausible but could, nevertheless, have proved true, as previous data show. The most appropriate comparison is with the corresponding condition in Experiment 6.2. There, the prior mean for unexpected claims was 4.29 on a scale from 0 to 10. Turning to reliability, the dependent measure was worded as ‘How reliable do you find [source]?’ This wording is rather abstract; more specific wordings could be used, asking people about trust, likability, and so on. While decomposing reliability in this way might uncover an effect, McCready’s (2014) account does not make predictions on this level. Neither design feature, then, seems to seriously compromise the results.

The data above may seem surprising in light of findings in other areas. The first of these is research into verbal probability expressions: modal adjectives, such as ‘impossible’, ‘possibly’, ‘likely’ and ‘certain’. Such terms feature in standardized lexicons expressing risks around climate change (Budescu, Broomell, & Por, 2009). One feature of verbal probability expressions is that they have directionality independently of their associated probability. That is, some words suggest event occurrence, while others suggest non-occurrence. For example, although ‘some possibility’ and ‘quite uncertain’ are given similar numerical interpretations, significantly more participants prefer an operation described as having ‘some possibility’ of success to one whose success is ‘quite uncertain’ (Teigen & Brun, 1999). Thanks to their directionality, verbal probability expressions have subtle effects on reputation. Teigen (1988) presented participants with predictions couched in verbal probability expressions with different directionality: one expert predicted a

RATIONALITY, PRAGMATICS, AND SOURCES

rise in crude oil prices by saying ‘It is possible that oil prices will reach \$20 in October’, another by saying ‘It is not quite certain that oil prices will reach \$20 in October’. Participants were then told that oil reached \$20. Even though participants thought the first expert had a lower probability in mind, they decided that he was more right than the second expert (for discussion, see Teigen & Brun, 1999).

The question naturally arises of whether directionality is essential to hedging. If so, then McCready’s (2014) account may be correct in spirit but wrong in detail: that is, hedging may be a crucial mechanism, but only hedging with propositional attitudes which have negative directionality. Certainly, the present studies used propositional attitudes with positive directionality. One test for directionality is to think of (or ask participants for) continuations: so, for example, to complete the sentence ‘It’s possible that X because...’, we would add a reason for occurrence; but to complete the sentence ‘I’m not completely certain that X because....’, we would add a reason for non-occurrence. Both ‘suspect’ and ‘am certain’ suggest occurrence. The shield hedge is more complicated: the sentence ‘I might be wrong, because...’ invites reasons for wrongness, but the phrase ‘I might be wrong, but...’ seems to override this directionality. Future work could manipulate directionality in hedging with propositional attitudes. However, if such hedging is to play a role in justifying cooperation, an account would also need to show that directionality is a semantic, rather than pragmatic, phenomenon. As we have already seen, it is circular reasoning to say that pragmatic inference is a precondition for pragmatics.

At first blush, there is also tension with a second area of existing research: work on plausible deniability and indirect speech (J. J. Lee & Pinker, 2010; Pinker, Nowak, & Lee, 2008). According to this research, a strategic speaker can choose to make indirect speech acts when it is unclear whether a context is one of cooperation

or conflict. In particular, a strategic speaker can enable a cooperative hearer to act favourably and prevent an uncooperative hearer from acting antagonistically.

Imagine, for example, a motorist stopped by a police officer and given a ticket. The motorist wishes to avoid a ticket by offering a bribe. A bribe would be accepted by a dishonest police officer, but would lead to arrest by an honest police officer. One option is to offer a bribe through an indirect speech act, such as ‘Gee, officer, is there some way we could take care of the ticket here?’ (Lee & Pinker, 2010; Pinker, Nowak, & Lee, 2008). This utterance would allow a dishonest officer to recognize and accept the bribe, but prevents an honest officer from having clear evidence for an arrest. There is experimental evidence that indirectness is sensitive to such pay-off structures (Lee & Pinker, 2010).

There is reason, however, to doubt that this tension is real. Plausible deniability may indeed be a significant motivator of indirect speech, but its effects do not depend on reputation management. Take the case of the motorist above, who uses an indirect speech act to offer a bribe. While the motorist may well have done enough to avoid being arrested, s/he is likely to suffer reputational damage. Indeed, such manoeuvring can be transparent and can itself cause reputational damage without undermining its ultimate strategic objective.

As we have seen, evidential language might yet allow reputation management but through a pragmatic mechanism: direct comparison with a set of alternative expressions, giving rise to a scalar implicature. But how might evidential language be interpreted without this kind of comparison or, indeed, if the proposed pragmatic mechanism does not exist? We can again draw a comparison with verbal probability expressions. Verbal probability expressions have uses beyond expressing probabilities. Bonnefon and Villejoubert (2006) showed that verbal probability

expressions can be used to convey bad news tactfully. They had participants read the sentence, ‘The doctor tells you, you will possibly suffer from insomnia [deafness] soon’ and elicited membership functions from participants.⁶¹ Participants understood ‘possibly’ as indicating higher probabilities with deafness than insomnia. Moreover, some 60% of participants indicated that the doctor was being tactful not uncertain. Bonnefon and Villejoubert argue that apparent severity biases can result from participants perceiving an utterance as a face-threatening act: that is, as an act that threatens someone’s desire for autonomy (their negative face) or their desire for connection with others (their positive face) (Brown & Levinson, 1987; Holtgraves, 2002). On Bonnefon and Villejoubert’s account, verbal probability expressions can serve as face-management devices: they allow a speaker to acknowledge, and lessen the impact of, a face-threatening act.

This politeness account can be extended to the present data. Participants may have understood at least some of the claims as advice: most plausibly, advice to use valium, buy an oven, bet on a horse, or attend a nightclub. Participants might take this advice to threaten their negative face: advice is not far removed from requests, offers, compliments, and so on, which are typically taken to threaten negative face (for discussion, see Holtgraves, 2002). For this suggestion to have traction, though, it needs empirical support. The simplest method would be a straightforward survey, in the manner of Bonnefon and Villejoubert (2006), asking participants why the speaker used the propositional attitudes in question. Although a polite source could

⁶¹ In membership-function studies, participants see a number line representing the probability range [0,1]. The line picks out numbers at even intervals across the range, different studies selecting different intervals. Participants are asked to rate how well each number corresponds to a verbal probability expression. These data allow a membership function to be calculated which can be interpreted as representing the meaning of the expression.

seem cooperative, helpful, and so on – hence, in some sense, reliable - they may nevertheless not appear reliable in the sense of being a source of true information. It could be instructive, therefore, to manipulate the social setting and, with it, the plausibility of face-management strategies.

7.5.1 Conclusions

The discussion above is not promising for McCready's (2014) account of hedging: two principal methods do not block downward revisions of source reliability. What impact do these findings have for pragmatic theory? Three possibilities present themselves. Firstly, reliability is truly a precondition for extended cooperation and pragmatics. In this case, when sources make unexpected claims and become increasingly unreliable, hearers should be less inclined to cooperate and draw pragmatic inferences. If sources wish to manage their reputation, they may have to take more drastic measures, such as presenting only messages they think their audience will expect, or doing substantial groundwork to change the audience's expectations, for instance, through extensive argumentation. Secondly, reliability is, indeed, a precondition for extended cooperation and pragmatics, but hearers are epistemically vigilant only some of the time. Unless hearers are given cause for concern, they may simply accept the information they receive. Thirdly, reliability is not a precondition for extended cooperation and pragmatics. Hearers may well need to adopt an initial stance of trust. But even when they come to judge a source highly unreliable, they commit the effort to understand the source's claims, withholding only their acceptance.

Each option above is plausible. It will take a larger research project to decide between the possibilities. Such a project would need to engage with contexts where pragmatics might proceed without trust and cooperation, such as legal cases or

RATIONALITY, PRAGMATICS, AND SOURCES

personal conflicts (J. Goodwin, 2001). For present purposes, though, the data and discussion serve to underline the close relationship between pragmatics and testimony, hence between pragmatics, sources and rationality.

8 General Discussion

This thesis has introduced a triangular scheme for the relationship between rationality, pragmatics, and sources. Throughout, it has argued for close connections between the components, and suggested that the study of any one component benefits from considering the others too. This chapter will recapitulate the main findings, will interpret them further, and will develop a more detailed account of the relationship between the components. In so doing, it will return to themes from Chapter 1, in particular to argumentation, which will serve as a test case for understanding the relationship between rationality and pragmatics.

8.1 The Experimental Data Recapitulated

8.1.1 Conditionals and Testimony

A large section of this thesis has looked at how we change our belief when we encounter a testimonial conditional: that is, a conditional uttered by a source of partial reliability. Testimonial conditionals show the close relationship between all three components; previous chapters have highlighted the link between pragmatics and sources, since the experiments focus on the role of source information in simple pragmatic contexts.

Chapters 2 to 4 reported a series of experiments focusing on belief change from the assertion of a conditional by single or multiple sources and by sources that were low or high in expertise. The data offered new insight into belief change from conditionals. Chapter 2 reported a set of experiments in which participants responded using point values. These data suggested that belief change is localized to the conditional probability. Chapter 3 reported experiments which replicated these effects, allowing participants to respond with intervals. These data suggested localized belief change in the form of an increase in the conditional probability and a

narrowing of the intervals. Throughout, the studies showed a diminishing return of multiple sources, which is to say that there were significant increases from null to single conditions but not from single to multiple conditions. Chapter 4 explored this issue further, and found suggestive – but qualified – evidence that multiple independent assertions were enough to cause a significant increase. A more robust finding in Chapter 4 was that prior belief mattered: experiments showed that the probability of the antecedent increased when its prior probability was low, and decreased somewhat or stayed flat if it was high; the same pattern was found for the probability of the consequent.

The data above have broad relevance for theories of the meaning of the conditional and, consequently, for theories of conditional reasoning. The data concern, that is, the interpretation of the premises in conditional reasoning tasks. The data are not straightforwardly accommodated by any individual theory. Chapters 2 and 3 suggested that belief in the antecedent did not change; Chapter 4 suggested that such belief could indeed shift, depending on the prior probability. But both of these points contradict the material-conditional and Kullback-Leibler theories of the conditional, according to which, on learning a conditional, the probability of the antecedent should always decrease. The data are weakly consistent with the suppositional theory of the conditional. The theory takes assertion to imply that the conditional probability is high, and can readily be extended to allow the conditional probability to be sensitive to source reliability (Stevenson & Over, 2001). However, the theory makes no predictions about the probability of the antecedent or the probability of the consequent. One way to extend the theory is to use Bayesian belief networks, graphical models which capture probabilistic relationships among random variables. These networks can be construed as representing the meaning of the

conditional: the meaning is the potential for belief change across the network. These networks can recreate the experimental data to some extent, but the relevant patterns depend on conditions which are not obviously met and whose realism is questionable. Chapter 4 tentatively suggested a pragmatic account, and argued for further, more psycholinguistic studies to ascertain what is semantic and what is pragmatic.

The data in Chapters 2 to 4 are a step towards developing normative testimonial models for conditionals. The empirical data from pragmatic contexts can help to constrain such models. For developing normative accounts, one problem has been a lack of consensus on the meaning of the conditional. Meaning has been approached obliquely by the present studies, which have considered the beliefs that a hearer is prepared to revise. More data is, of course, necessary. But, knowing about these beliefs, we can consider whether it is rational to revise them and, if so, in which direction.

8.1.2 Utility Conditionals

Chapter 5 also studied conditionals, but this time from the point of goal framing. Its findings bear on the link between pragmatics and rationality. The experiments in Chapter 5 treated goal frames as arguments from consequences; their strength was taken to depend on the utilities and the probabilities. Experiments tested the role of utilities. The data supported such a role, though in an unanticipated way. Experiment 5.1 showed that, contrary to initial predictions, the positive frame was more persuasive, and this resulted in part from an unexpected assignment of utilities. Many participants assigned positive utilities to the negated actions ('taking mildly unpleasant medicine' or 'undergoing painful surgery'); with this utility assignment, the negative frames are ineffective arguments, because the positive utility of the antecedent balances out the negative utility of the consequent. Experiment 5.2

RATIONALITY, PRAGMATICS, AND SOURCES

replicated these data with different ways of measuring convincingness and utility.

There was further evidence that negated unpleasant actions were assigned positive utility. Negated positive consequents were, conversely, assigned disutility.

Experiment 5.3 addressed the ambiguity in the negated consequents: that a risk not decreasing could mean either the risk increasing or not changing. The experiment compared the rewordings ‘the risk will increase’, ‘the risk will stay the same’, and ‘the risk will not change.’ The data suggested that participants understood ‘not decreasing’ to mean increasing, further casting doubt on the general convincingness of negative frames.

Chapter 5 offered various arguments for the closeness of pragmatics and rationality, both explicit and implicit. Firstly, the chapter stressed the closeness of speech-act theory and persuasion; persuasion fits neatly into the category of perlocutionary effects, an issue we will return to below. Secondly, the chapter showed that utilities are crucial to theories of both pragmatics and argumentation. Thirdly, the chapter derived initial predictions from the theory of utility conditionals, which applies to both pragmatics and argumentation. Fourthly, and most substantively, the data were interpreted as showing a complex interaction between pragmatics and utilities. The crucial phenomenon was negation. Negation targeted the action (antecedent) or consequence (consequent), but also placed focus on the fact that expectations had been violated: it was in some way expected that the hearer should take the action. A subtler explanation applies to the consequence: in the experiments, no clear expectations arose for the ‘not change’ wording; hence, the utility was not reliably different from ‘stay the same’. This negation-and-expectations account makes sense of the utility assignments. However, the explanation does not sit well with existing studies which suggest that utilities are

used to generate pragmatic inferences, not *vice versa*. Chapter 5 suggested a complex process whereby basic pragmatic processes prompt utility assignments, reasoning on which generates subsequent pragmatic inferences.

8.1.3 Testimony and Source Reliability

Chapter 6 explored how people respond to information from partially reliable sources. As such, it focused on the link between sources and rationality. The chapter found evidence for epistemic vigilance: people were sensitive to both the content of claims and the reliability of the source. These types of information also interacted: expected claims increased reliability, and unexpected claims decreased it; reliable sources increased belief, and unreliable sources decreased it. This last effect was the most striking: sources could be anti-reliable. This epistemic vigilance is intuitively appealing. The data also bear on the relationship between sources and pragmatics, by emphasizing the difference between comprehension and acceptance, and by raising questions about how such vigilance impacts on extended cooperation and pragmatics, if pragmatics requires trust to proceed.

8.1.4 Evidential Language

Finally, Chapter 7 explored this tension between epistemic vigilance and pragmatics. The experiments explored McCready's (2014) theory that speakers manage their reputations with hedging, so that they can preserve the cooperation necessary for pragmatics. The experiments found no evidence that hedging with propositional attitudes or with shield hedging provided any protection for sources making unexpected claims. Sources making such claims suffered robust decreases in their source reliability. These data cast doubt on a role for hedging in justifying pragmatic inferences. Chapter 7 discussed various possibilities: that speakers use more dramatic means to protect their reputations, such as avoiding making

unexpected claims; that hearers are only sometimes epistemically vigilant; and that pragmatics does not, in fact, rely on trust and cooperation.

Having summarized the experimental data, we can turn to common themes across the studies, and to a deeper analysis of the links between rationality, pragmatics, and sources.

8.2 Redefining the relationships

In the course of this thesis, we have considered how to define communication⁶², and how the definition may help to understand the relationships among the triangular scheme of pragmatics, rationality and sources. This definition serves as a springboard for the discussion below. Let us return, then, to the definition of communication from Relevance Theory (Sperber & Wilson, 1995):

Informative Intention: to inform the audience of something.

Communicative Intention: to inform the audience of one's informative intention.

To recapitulate the discussion in Chapter 6, communication occurs if either intention is fulfilled. Communication occurs if only the informative intention is fulfilled, through presentation of direct evidence. For instance, the speaker can point to a state of the world or provide a logical argument; the audience can, in principle, change their belief in the intended way without recognizing the intention. Communication also occurs if only the communicative intention is fulfilled, through the clear signalling of an informative intention which is recognized by the audience but not

⁶² Since, on this definition, communication is fundamentally pragmatic, I slip, in this section, between talking about communication and talking about pragmatics. Treating communication in this way is not wholly unproblematic. It seems to follow, for instance, that pragmatically naïve young children – or those with pragmatic dysfunction – are not capable of genuine communication. I set aside this topic here. For discussion, and a possible solution, see Breheny (2006).

fulfilled. For instance, an audience understands an obvious lie but does not believe it. Finally, communication occurs if both intentions are fulfilled: in this case, the audience recognizes the speaker's intention and also changes their belief in the intended way.

With this definition, it becomes clearer why communication and more traditional rational behaviour are so closely related. We can see this most clearly with two examples: the relationships between communication and testimony and between communication and argumentation. Take, first, the relationship between communication and testimony. If only the informative intention is fulfilled, then communication occurs but of a type that does not fit the definition of testimony. As long as the communicative intention is fulfilled, then there is testimony; and if both intentions are fulfilled then there is successful testimony. Successful testimony is pragmatics (communication) with epistemic authority (for discussion, see Sperber et al., 2010). To put matters differently, there is a distinction between comprehending (fulfilling only the communicative intention) and accepting (fulfilling both intentions). At least in principle, the two diverge and it is possible to treat testimony critically.

How does this distinction between comprehension and acceptance map onto the empirical data in this thesis? There is indirect evidence of this distinction in Chapters 2 to 4: ratings were significantly lower for inexpert than expert sources, and were not at ceiling for expert sources. In other words, there was presumably a gap between comprehension and belief. More direct evidence would result from adapting the paradigm to ask about the speaker's beliefs or their intended meaning, that is, what the speaker wanted the hearer to believe. But there is more direct evidence of the distinction in Chapter 6, where participants clearly understood

RATIONALITY, PRAGMATICS, AND SOURCES

claims without accepting them; indeed, they understood claims, and revised their belief *away* from the claimed state of affairs. The evidence – both indirect and direct – suggests critical faculties which are distinct from pragmatics.

At first blush, this apparent scepticism does not sit well with other data on belief fixation. This tension needs resolving. As we have seen throughout this thesis, experimental work suggests that people conform to a Spinozan model of belief fixation: when a hearer entertains a proposition, by default they passively accept it; they can then optionally either actively endorse or actively reject the proposition (D. T. Gilbert et al., 1990, 1993; Mandelbaum, 2014). That is, it takes extra work to assess a proposition critically; doing so is optional and the process is prone to disruption. Supporting evidence comes from Gilbert and et al.'s (1990, 1993) experiments in which cognitive load caused false information to be misremembered as true. Other evidence comes from belief-perseverance tasks and response biases in personality questionnaires (Mandelbaum, 2014). There are, yet, defenders of more objective belief fixation (Hasson et al., 2005), and it is clearly possible for people to treat claims sceptically (see, for example, Chapters 2 to 4 and 6). However, the possibility arises that, in early processing, comprehension and acceptance are one and the same, and that acceptance simply collapses into comprehension. More data is needed to explore this possibility for the data in Chapters 2 to 4 and 6, as there are clear implications for processing measures: for instance, sceptical responses should take longer, and should be less likely under cognitive load.

The definition above also offers insight into the relationship between pragmatics and argumentation. Communication, as we have seen, can be successful if only the informative intention is fulfilled, by providing direct evidence that something is the case. For example, someone could show an audience a photograph

of a particular person attending a party and thereby inform the audience of the fact. But, for present purposes, a more interesting case is an argument: a speaker can lay out an argument to inform a hearer of a proposition, the conclusion. The hearer can accept the proposition without recognizing the communicative intention behind it. Of course, the hearer might also recognize this communicative intention but need not. In many cases, it might prove practically impossible to show whether a proposition is accepted because of the argument or because of the communicative intention.

This definition of communication allows us to understand the relationship between pragmatics and rationality somewhat better. But still greater understanding will come from a deeper analysis of pragmatics and of a behaviour which seems uncontroversially about rationality: namely argumentation.

8.2.1 Pragmatic intrusion into argumentation

The study of argumentation shows intimate connections with pragmatics. The connection can be seen pre-theoretically in the following example conversation from Collins and Hahn (2016), patterned after the analyses in Clark (2013):

A: John looks grumpy.

B: He hasn't had breakfast, so he is starving.

(3) [John] hasn't had breakfast, so [John] is starving.

(4) [John] hasn't had breakfast [today], so [John] is starving.

(5) [John] hasn't had breakfast [today], so [John] is [very hungry].

(6) [B believes that][John] hasn't had breakfast [today], so [John] is [very hungry].

(7) [B believes that John looks grumpy because (or so B believes)][John] hasn't had breakfast [today], so [John] is [very hungry].

RATIONALITY, PRAGMATICS, AND SOURCES

This apparently simple conversation is underpinned by a series of inferences to an implicature, sentence (7). Something like these inferences presumably occurs in a hearer's head when understanding implicatures. These inferences look much like a rational argument.

The connection between pragmatics and argumentation has been acknowledged for some time in argumentation theory. As we saw in Chapter 1, argumentation theory took a social turn, starting with the work of Toulmin (1958), such that argumentation research came to focus on how people use propositions to achieve a goal. This social turn paralleled the turn, in philosophy of language, towards the use of language, which gave birth to natural-language pragmatics. In argumentation theory, the social turn comprised, in large part, acknowledging that argumentation takes place in dialogues. More recent argumentation theory has fleshed out the different types of dialogue and their different goals. Table 8.1 shows one possible taxonomy of dialogues.

Table 8.1. Dialogue types according to Walton (2008, p. 8)

Type of dialogue	Initial situation	Participant's goal	Goal of dialogue
Persuasion	Conflict of opinions	Persuade other party	Resolve/clarify issue
Inquiry	Need for proof	Find, verify evidence	Prove (disprove) hypothesis
Negotiation	Conflict of interests	Get what you most want	Reasonable settlement for both
Information-seeking	Need information	Acquire/give new information	Exchange information
Deliberation	Dilemma/practical choice	Coordinate goals and actions	Decide best available action
Eristic	Personal conflict	Verbally hit out at opponent	Reveal deeper basis of conflict

RATIONALITY, PRAGMATICS, AND SOURCES

In an argument, the first challenge, then, is to determine the type of dialogue. This is a non-trivial task, relying on various pragmatic and otherwise mentalistic resources. The task is not adequately explained in the argumentation literature, although dialogue types bear at least a superficial resemblance to scripts or event schemas (Groome, 2013; Schank & Abelson, 1977), suggesting an avenue for future research. However the task is achieved, identifying the dialogue type is a crucial step for a normative analysis: a key standard is the relevance of an argument to its dialogical goals; and many errors stem from mistaking the type of dialogue or switching between types (Walton, 2008).

Argumentation theory places considerable emphasis on norms. These norms often have a decidedly pragmatic flavour. Take Walton's (2008) proposals. On his account, there are norms for the types of speech act allowed; norms for appropriate turn taking in the dialogue; norms for the links between speech acts and commitments in the argument; and norms for determining, where appropriate, who has won or lost and how the argument should end. Walton also emphasizes overarching rules. Participants in arguments must be relevant. They must also be cooperative, answering questions, making and accepting commitments that follow from their stated positions, and fairly characterizing their opponent's position. And they must provide just enough information to convince, and avoid giving redundant information. Strikingly, this paragraph would be perfectly coherent if the word 'argument' were replaced with 'conversation' (excepting wins and losses).

The connection between pragmatics and argumentation is even more explicit in the pragma-dialectic approach to argumentation. Here, argumentation is presented as 'normative pragmatics' (Eemeren & Garssen, 2013). One way of understanding this term is that studying arguments is a lot like studying pragmatics. To interpret an

RATIONALITY, PRAGMATICS, AND SOURCES

argument, a researcher must select only the information relevant to the argument; must make explicit any necessary implicit information (e.g. unexpressed premises); must re-order the information in a logical form; and must represent the argument in a uniform notation (Eemeren & Garssen, 2013). In pragmatics, researchers do not typically delete information, but they follow the remaining steps: witness, for example, the analysis of the conversation above.

Although these similarities are suggestive, a more substantive case lies in the pragma-dialectic account of argumentation itself (see, e.g., Eemeren & Grootendorst, 1992, 1995, 2004). Pragma-dialectics is closely related to Walton's (2008) framework, but emphasizes the pragmatic nature of argumentation still further. Like Walton, pragma-dialecticians take argumentation to be goal-based: the acceptability of an argument is determined by its capacity to achieve its goal. Again like Walton, pragma-dialecticians take argument to be an interaction between at least two individuals, and to be about resolving differences of opinion between these individuals. But on one point they are more explicit: argumentation is a complex speech act (see, e.g., Eemeren & Garssen, 2013). This speech-act account requires an extension to standard speech-act theory. The standard theory allows for five types of speech act: directives, which are attempts to get a hearer to perform an action (e.g. requests); assertives, which commit the speaker to the truth of a proposition (e.g. assertions); commissives, which commit the speaker to an action (e.g. promises); declaratives, which are attempts to change the state of the world through some institutional power (e.g. performing a marriage); and expressives, which express a psychological state (e.g. thanking or complaining) (Holtgraves, 2002; Searle, 1969). An argumentative speech-act theory has to allow speech acts to have dual roles: the standard roles above and their role in the 'higher textual level', which is to say their

RATIONALITY, PRAGMATICS, AND SOURCES

role in the argument (Eemeren & Garssen, 2013). An assertive, for instance, lets the speaker express their standpoints in the argument; a declarative allows the speaker to challenge someone to defend their argument; and a commissive allows the speaker to accept or reject a proposed standpoint or argument (Eemeren & Garssen, 2013).

Pragma-dialectics echoes Walton's (2008) approach to norms, but provides a more extensive account. The theory integrates pragmatic principles into argumentation: in Gricean style, arguers are required to be clear, honest, efficient, and to the point (Eemeren & Grootendorst, 1992). But pragma-dialectics offers additional norms. For instance,

Rule 1 (Freedom Rule): Discussants may not prevent each other from advancing standpoints or from calling standpoints into question.

Rule 2 (Burden of Proof Rule): Discussants who advance a standpoint may not refuse to defend this standpoint when requested to do so.

(Eemeren et al., 2009, pp. 21–22)

These additional norms are largely procedural: that is, they regulate behaviour without engaging with the actual content of an argument. The norms can account for some informal fallacies. For instance, Rule 1 prevents threats of force (*argumentum ad baculum*), appeals to compassion (*argumentum ad misericordiam*) or personal attack (*argumentum ad hominem*) (Eemeren & Grootendorst, 1995). Pragma-dialectic rules are 'not metaphysically based, but rest on their suitability for doing the job for which they are designed' (Eemeren & Garssen, 2013, p. 518). In other words, rules are valid because they work: in general, they help to resolve disputes. This is a somewhat weaker rational justification than is often seen in the psychology of rationality.

Pragma-dialectics has one final component which illustrates the pragmatic – or, at least, mentalistic – character of argumentation: strategic manoeuvring. On this view, arguers maintain a delicate balance between the reasonableness and effectiveness of an argument. Arguers can strategically select topics, adapt their arguments to their audience, and adjust their presentation of the argument to minimize the ‘disagreement space’ (Eemeren & Houtlosser, 1999). This manoeuvring allows arguers to smuggle into a particular argument moves which are generally unreasonable. For instance, arguers can disguise an *ad hominem* argument so that it appears considerably more reasonable (Eemeren, Garssen, & Meuffels, 2012). Construed broadly, strategic manoeuvring corresponds to presentational devices shown to be effective in argumentation, such as explicit marking of conclusions, identifying sources, and use of figurative language (O’Keefe, 2012).

The discussion above shows considerable pragmatic intrusion into argumentation theory. The theories in question, however, arguably provide only a shallow normative account, a shallowness which could – but, in practice, does not – limit the pragmatic intrusion into rationality. The main challenge to these theories is that their norms do not have a deep philosophical basis. They contrast starkly with, say, the norms of logic or probability theory. If the theories do not engage with the actual content of an argument, they also encounter substantial problems. We can readily imagine two arguers who follow procedural rules to the letter: in such a case, there may be no clear way to decide which argument is better (Hahn & Hornikx, 2016). It is preferable for a theory of argumentation to deal with the content, as well as the procedure, of an argument. But, as the following section will show, there are links between such content-based theories and pragmatics. These links are illustrated, below, by considering the influence of content-based norms on pragmatic reasoning.

8.2.2 Argument in pragmatics

Pragmatic theories have made concrete suggestions about cognitive mechanisms underpinning pragmatics, be they refinements or replacements of Gricean maxims (Horn, 1984, 1989, 2004; Levinson, 2000; Sperber & Wilson, 1995). Arguably, though, ‘the crucial step between what is said and what is meant has never been described’ (Macagno & Walton, 2013, p. 204). To put the point more moderately, pragmatic theories have tended to stop short of mentioning specific argument forms, preferring abstract statements about reasoning being defeasible (e.g. Sperber & Wilson, 1995). Pragmatic theories can gain in specificity by focusing on specific forms of argumentation.

Macagno and Walton (2013) argue that implicature should be treated as dialectical argumentation, much in the manner of the argumentation theories discussed above. Their account has the following basis. Firstly, Gricean pragmatics should be weakened, because regulatory principles are context-dependent. Take, for instance, the following court-room exchange, analysed in Goodwin (J. Goodwin, 2001):

Q: Do you have any bank accounts in Swiss banks, Mr Bronston?

A: No, sir.

Q: Have you ever?

A: The company had an account there for about six months, in Zurich.

In normal conversation, the final answer implies that the speaker had never personally had a Swiss bank account. In fact, the speaker *had* had a Swiss bank account, but he was found by the court to have replied truthfully. Such evasions are not prohibited in court-room pragmatics (J. Goodwin, 2001; Macagno & Walton, 2013). Pragmatic principles are, on this view, relative to dialogues – to the very same

dialogue types in Table 8.1. Indeed, ‘a type of dialogue...can be considered as a normative framework in which there is an exchange of arguments between two speech partners reasoning together in turn-taking sequences aimed at moving toward the fulfilment of a collective conversational goal’ (Macagno & Walton, 2013, p. 206).

Dialogues, then, are key to this argumentative account of pragmatics, but how do they influence utterance interpretation? Dialogue types provide overarching goals which ‘assign a role to each dialectical move’ (Macagno & Walton, 2013, p. 207) and make available presumptions about the behaviour of dialogue participants – presumptions about what behaviours are acceptable in a given dialogue. To borrow Macagno and Walton’s example, consider an utterance such as ‘I like eating at Gino’s’. The dialogue type directly influences its interpretation. If the dialogue is a deliberation about where to eat, then the utterance is a proposal and implicit argument (‘We should go to Gino’s, because I like it’). If the dialogue is an information-seeking dialogue about local restaurants, then the utterance is a simple personal assessment. If the dialogue is eristic (a personal argument), then the utterance could be a statement of personal offence (‘Only idiots like eating at Gino’s’ – *I like eating at Gino’s; are you calling me an idiot?*). Here, interpretation is guided by presumptions about behaviour; whenever there is a clash between competing presumptions, dialogue participants must reason their way to an explanation of the clash.

So far, the framework is as abstract as other pragmatic accounts. What makes the account more specific is its calling on argumentation schemes. Argumentation schemes are patterns of inference which are typically associated with a set of critical questions for assessing the strength of particular arguments (see, e.g., Walton et al., 2008). Macagno and Walton (2013) identify a range of schemes that are particularly

RATIONALITY, PRAGMATICS, AND SOURCES

relevant, arguing that the schemes are invoked when there is a clash of presumptions.

Take, for instance, their example:

Bert: Do you like ice cream?

Ernie: Is the Pope Catholic?

Bert's question sets up the presumption that this an information-seeking dialogue, and that Ernie will supply information in return. Ernie, however, responds with another question. One possibility is that he is being unreasonable, but this conflicts with the presumption that speakers usually are reasonable. More likely, then, is the possibility that Ernie's question is meant as a reply. Here, Bert can assume that Ernie's question is similar to his own, and can invoke the scheme for arguments from analogy (as reported by Macagno & Walton):

MAJOR PREMISE: Generally, case C_1 is similar to case C_2 .

MINOR PREMISE: Proposition A is true (false) in case C_1 .

CONCLUSION: Proposition A is true (false) in case C_2 .

Ernie's question has an obvious (affirmative) answer. Bert may reason, then, that his own question has an obvious (affirmative) answer.

Even simple exchanges can invoke multiple argumentation schemes. Take, for instance, a second example from Macagno and Walton (2013).

A: I am out of petrol.

B: There is a garage around the corner.

Speaker A is sharing information, but seems to be violating a presumption of information-sharing dialogues: that B is interested in this information. This leads to a reinterpretation of the utterance using an argumentation scheme: the appeal to pity (here, as cited by Macagno and Walton):

Individual x is in distress (is suffering).

RATIONALITY, PRAGMATICS, AND SOURCES

If individual y brings about A , it will relieve or help to relieve this distress.

Therefore, y ought to bring about A .

In other words, speaker A is making an appeal to pity – is asking speaker B to help him/her in some way. The presumed response to a request for help is an action.

Speaker B does not respond with an action, but rather with information. Speaker A , however, can interpret B 's response using an argumentation scheme for practical reasoning (again, as cited by Macagno and Walton):

My goal is to bring about A (Goal Premise).

I reasonably consider on the given information that bringing about at least one of $[B_0, B_1, \dots B_n]$ is necessary to bring about A (Alternatives Premise).

I have selected one member B_i as an acceptable, or the most acceptable, necessary condition for A (Selection Premise).

Nothing unchangeable prevents me from bringing about B_i as far as I know (Practicality Premise).

Bringing about A is more acceptable to me than not bringing about B_i (Side Effects Premise).

Therefore, it is required that I bring about B_i (Conclusion).

Speaker B , then, is making the implicit argument that Speaker A should get petrol at the garage around the corner.

Macagno and Walton (2013) offer forms of argument that are more specific than those usually discussed in the pragmatics literature. But there are problems with these argumentation schemes which make them questionable as either normative or descriptive accounts. Before considering these problems, it will be useful to see how argumentation schemes are supposed to work. Consider a further example from

RATIONALITY, PRAGMATICS, AND SOURCES

Macagno and Walton and the corresponding argumentation scheme, the Argument from Sign (Walton, Macagno, & Reed, 2008, p. 329).

A: What on earth has happened to the roast beef?

B: The dog looks very happy

Implicit Conclusion: The dog has eaten the roast beef.

Argument from Sign

A (a finding) is true in this situation.

B is generally indicated as true when its sign, A, is true

B is true in this situation

Critical Questions

CQ1: What is the strength of the correlation of the sign with the event signified?

CQ2: Are there other events that would more reliably account for the sign?

Here, the sign is the dog's happy expression: the critical questions are how strongly the dog's happy expression correlates with eating roast beef, and whether there are alternative events which would better explain the dog's happy expression. The answers to the critical questions determine the strength of the argument.

This argumentative account of pragmatics inherits the difficulties of the general approach to argumentation schemes, surveyed by Hahn and Hornikx (2016). Firstly, argumentation schemes proliferate: theories differ in the number they include, and principled choices do not seem to have been made about the inclusion and discreteness of the schemes. Secondly, it is not clear what normative status the critical questions have. It is unclear, for instance, what constitutes a successful answer and whether successful answers oblige the audience to accept the argument (Hahn & Hornikx, 2016; Walton et al., 2008). Relatedly, there is often no clear

formal link between the argumentation scheme and the critical questions; their link is only intuitive (Hahn & Hornikx, 2016). While the questions could, in principle, be added to the arguments as premises, the underlying logic can only be a defeasible *modus ponens*, an argument form which inherits all the defects of classical *modus ponens*, such as licensing highly counter-intuitive arguments in which there is no real relationship between the premises and conclusion (Hahn & Hornikx, 2016).

The defects in argument schemes naturally raise the question of whether there are suitable alternatives which could be applied to pragmatics. One such alternative is Bayesian Argumentation (Hahn & Oaksford, 2007). On this approach, argumentation is treated as probabilistic inference: at base, it is Bayesian belief revision. Bayesian Argumentation, therefore, has a firm normative foundation. Arguments are measured by two criteria: their strength, that is, the degree of belief in the conclusion; and their force, that is, the amount of belief change brought about. This Bayesian approach has led to a reassessment of numerous forms of informal argument, and the conclusion that many supposed fallacies are arguments which can, in the right circumstances, be strong. There are Bayesian accounts of the argument from ignorance (Hahn & Oaksford, 2007), circular reasoning (Hahn & Oaksford, 2007), the slippery slope argument (Hahn & Oaksford, 2007; Corner, Hahn, & Oaksford, 2011), and the argument from expert opinion (Hahn et al., 2012; Harris et al., 2016). Hahn and Hornikx (2016) also offer re-analyses of the argument from sign and the argument from popular opinion. Bayesian Argumentation, in sum, provides a promising, if largely unexplored (though see Harris et al., 2013), approach to pragmatics.

As we have seen, then, pragmatic explanations can be made more concrete through argumentation schemes, which can be re-expressed probabilistically to

provide a sure normative foundation. But there is arguably a still deeper connection between argumentation and pragmatics: they are both underpinned by relevance (for discussion, see Collins & Hahn 2016). Consider, first, argumentation. The quality of an argument is, in large part, determined by its relevance to some goal (Walton, 2003; Walton & Macagno, 2016). That is, a good argument is one which is capable of changing beliefs towards that goal (Walton & Macagno, 2016). Walton and Macagno give the following example (Walton & Macagno, 2016, p. 526):

Your friend says that Brand X coffee tastes better than Brand Y. Apparently she is ignoring the fact that Brand X is made by a company that also made a product that was responsible for thousands of deaths of children in undeveloped countries. Therefore, your friend is mistaken.

The relevance of this argument depends on the goal. If the friend in question is merely expressing her personal opinion on the taste, then the counterargument is irrelevant: it is not capable of changing beliefs about the coffee's flavour. A relevant counterargument might be to mention a new, improved version of Brand Y which might change the friend's view. However, if the goal is to decide on a coffee to buy, then the counterargument is relevant: it is an ethical argument against buying Brand X. In more realistic, more complex argumentative discourses, there may be many interlinked arguments, each with local goals. Important, in such cases, is the ultimate goal, or *stasis*, of the dialogue and the capacity of each argument to shift belief towards achieving that goal (Walton and Macagno, 2016).

The idea of a stasis has a direct analogue in contemporary semantics and pragmatics: the Question Under Discussion, or 'QUD' (Roberts, 1996). On this account, any discourse raises a set of questions, implicit or explicit, to be addressed. These questions are typically viewed as sets of alternatives or competing

descriptions of the world (Clifton & Frazier, 2012). The utterances that make up the discourse can be tied together by considering how they relate to the overall QUD(s). More philosophical treatments suggest that a discourse can raise an extremely large number of QUDs, and that these questions are sometimes only identified after the fact (Clifton & Frazier, 2012). In contrast, Clifton and Frazier (2012) propose a more computationally tractable account, focusing on a limited number of QUDs. On their account, QUDs are raised by explicit questions; by assertions with non-actuality implicatures ('He should have left work by now' - so has he?); by focus ('TOM kissed Mary' - who were the other potential kissers?); by disjunctions; and by indefinites ('I'm going to bet on a horse' - which one?). Clifton and Frazier used eye tracking to find evidence that people expect utterances to address QUDs, and that utterances which fail to do so disrupt processing.

For both concepts – stasis and QUD – relevance is a central issue. Macagno and Walton (2016) provide an analysis, for stasis, in terms of 'profiles of dialogue': graphs which show the structure of complex arguments, and can be used to judge whether individual components are relevant to the overall goal. Here, as with argumentation schemes, we can give a stronger account by invoking probabilities. Probabilities offer various ways of measuring relevance. One such we have seen in Chapter 5: the delta-P rule. Another is the difference rule: on this rule, an argument is relevant if it increases belief in the conclusion; hence, if $P(\text{Conclusion}|\text{Evidence}) > P(\text{Conclusion})$. This rule links smoothly with Bayesian belief networks. Since these graphical devices have well-known formal properties, they are invaluable in representing complex relations among random variables, and help to simplify calculations. Bayesian belief networks have been applied to simpler arguments and testimony, and could readily be extended to complex dialectical arguments. If these

methods can be used to represent all the evidence in a complex court case (Kadane & Schum, 1996), they can surely be applied to a dialectical argument. These methods also allow a dynamic notion of relevance: interventions in the network alter patterns of probabilistic dependence, thanks to the axioms of conditional independence (Korb & Nicholson, 2011). For instance, take Korb and Nicholson's (2011, p. 39) example of a model in which A (smoking) causes B (cancer) which causes C (shortness of breath). To quote:

If we don't know whether [a] woman has cancer, but we do find out that she is a smoker, that would increase our belief both that she has cancer and that she suffers from shortness of breath. However, if we already knew she had cancer, then her smoking wouldn't make any difference to the probability of [shortness of breath]. That is, [shortness of breath] is conditionally independent of being a smoker given the patient has cancer.

This flexibility seems highly advantageous.

Relevance is also an important notion in pragmatics, further emphasizing the deep connection with argumentation. Relevance is typically given a rather different treatment from that above: Relevance Theory, for instance, treats relevance as a trade-off between cognitive effects and processing effort (Sperber & Wilson, 1995). Cognitive effects are the strengthening or weakening of assumptions, and the derivation or cancellation of conclusions, and so on. In short, these are phenomena that could easily be cashed out in probabilistic terms. But it remains to be seen whether Relevance Theory can be re-expressed using Bayesian belief networks. As we have seen, though, there is certainly a sense in which discourse hangs together. Individual discourse moves seem relatable to an ultimate conversational goal in much the same way that individual arguments are relatable to an ultimate

RATIONALITY, PRAGMATICS, AND SOURCES

argumentative goal. Both types of goal will typically be achieved by modifying belief. And both types of move – conversational and argumentative – can vary in their instrumental rationality: that is, in how likely they are to bring about the overarching goal.

Much as there has been a blurring of the boundaries between pragmatics and argumentation, it is important to recognize the distinctions between the two abilities. We can see these distinctions at work in the following example argument: ‘If we allow gay marriage, then people will want to marry their pets’. Even without context, this is clearly a consequentialist argument. The relevant parameters are the utilities of the antecedent and consequent and the conditional probability. All else equal, the argument is likely to be understood as an argument against gay marriage: the (presumed) consensus is that interspecies marriage has disutility. This assignment of disutility is enough to interpret the conditional as an argument against gay marriage. But an audience might go further and conclude that the arguer considers gay marriage to have intrinsic disutility: the arguer is likely to be the kind of person who opposes gay marriage irrespective of its effects on interspecies marriage. The audience might also conclude that the arguer suspects that the audience considers gay marriage to have intrinsic utility: why make the argument if not to change the audience’s mind? The assertion of the conditional also implies that the arguer considers the conditional probability to be high. In context, there are potentially many other cues to these assignments: for instance, the arguer may have a set role in a deliberation, or may show emotion through intonation or facial expressions. But none of this is enough to make the argument persuasive. For persuasiveness, it is the audience’s utility and probability assignments that are key.

RATIONALITY, PRAGMATICS, AND SOURCES

The preceding discussion has sketched the following picture: pragmatics and testimony, and pragmatics and argumentation, use much the same patterns of inference, albeit it to slightly different ends. The question naturally arises of whether the different abilities rest on the same mechanisms. Here, a variety of methods suggest themselves. Clinical research suggests that a variety of developmental, psychiatric and acquired disorders disrupt pragmatics, including, for instance, autistic spectrum disorder, Williams syndrome, motor neuron disease, multiple sclerosis, schizophrenia, traumatic brain injuries, and strokes (Cummings, 2017). If pragmatics and argumentation rely on the same mechanisms, then pragmatic dysfunction should associate with argumentation dysfunction. A major challenge, however, would be to isolate pragmatic dysfunction from other linguistic and cognitive deficits which occur in many of the above disorders (Cummings, 2017). A related, but more straightforward, approach would be search for individual differences within the normal population. There is scant individual-differences research in either pragmatics or argumentation, but there are numerous plausible individual-difference variables, such as the communication subscale of the Autism Quotient questionnaire (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; Nieuwland, Ditman, & Kuperberg, 2010) or scales such as Need For Cognition (Cacioppo & Petty, 1982), Need for Affect (Maio & Esses, 2001), Need for Closure (Webster & Kruglanski, 1994), and so on. Such research would offer valuable insight into the abilities themselves and their relationship, and help to address fundamental questions such as the whether pragmatics is a modular process (Cummings, 2017; Sperber & Wilson, 2002).

8.2.3 Pragmatics as rational social action

The triangular scheme sits well with an emerging body of research which treats pragmatics as rational social action. This research uses formal mathematical tools to model language use in context, and increasingly leads to predictive psychological accounts of linguistic phenomena. Much of this research draws on game theory. Game-theoretic models of conflict and cooperation have abundant applications to meaning. Take, for instance, a simple zero-sum game of perfect information: that is, a game in which pay-offs sum to zero - so that in, for example, a two player game, one player's gain is offset by another's loss - and in which each player knows all previous events. Imagine a two-player version: one player is the verifier, who wins if they can demonstrate that a proposition holds; the other is the falsifier, who wins if they can demonstrate that a proposition does not hold. Assume, further, that we have a simple tree-adjoining grammar for (a fragment of) English, and that we have a simple model which maps English expressions onto entities in the model: real-world entities map onto constants; transitive verbs map onto sets; and intransitive verbs map onto ordered pairs. English propositions hold if they have counterparts in the model. Together, the model and game can be used to reconstruct the basics of truth-conditional semantics, building up from atomic sentences to negation, other connectives, quantifiers, pronouns, and scope phenomena (R. Clark, 2011).

Impressive though this feat may be, it relies on tenuous assumptions: for instance, that people have perfect information of previous events, that expressions are unambiguous, and that linguistic behaviour relies on conflict (R. Clark, 2011). We can reject these assumptions and prefer alternative models which better capture pragmatics. For instance, we can turn to games of partial information (R. Clark,

2011; Parikh, 2001). These are games in which one player, a speaker, selects a signal and another player, a hearer, must select an interpretation of the signal. Signals give rise, probabilistically, to different interpretations. These probabilities accommodate context. For instance, the word 'pen' is ambiguous between a writing instrument and an enclosure for animals. These interpretations have different probabilities. But the probabilities will change depending on context: at an academic conference, both interpretations are possible, but the writing instrument is more likely; at a farm, both interpretations are, again, possible, but the animal enclosure is considerably more likely than at the conference. Interpretations have utilities: the correct (intended) interpretation has utility for both speaker and hearer; the incorrect (unintended) interpretation has disutility for both speaker and hearer. These utilities can also depend on the signal chosen: more complex signals have lower utilities to encode a preference for economy. Speaker and hearer alike make strategic choices to maximize their expected utility. Games of partial information can model the resolution of lexical ambiguity, responses to garden-path sentences, reference assignment for pronouns, politeness behaviours, and typicality effects (R. Clark, 2011). To some extent, they can also model individual differences, allowing for players who prefer rewards and risk misunderstandings to achieve them, and for players who prefer certainty and risk expending effort unnecessarily (R. Clark, 2011).

A more recent alternative arises from an attempt to address problems with the use of standard game theory to explain decision making in social contexts: the theory of Virtual Bargaining (Misyak & Chater, 2014; Misyak, Melkonyan, Zeitoun, & Chater, 2014). Typically, standard game-theoretic explanations are based on Nash equilibria (Misyak & Chater, 2014; Misyak et al., 2014): an equilibrium arises when players have each selected a strategy, and no player will benefit from unilaterally

changing strategy. In some games, there is more than one equilibrium; in others, at least some real-world players systematically choose non-equilibria. For instance, some players of the Prisoner's Dilemma will choose to cooperate, even though the only equilibrium is for both players to defect (Misyak & Chater, 2014). A theory of social decision making needs to predict choices in both cases. Virtual Bargaining holds that decision makers will choose the option they would have chosen in explicit negotiation with other players. Although this research is primarily theoretical thus far, initial data are consistent with the theory (Misyak & Chater, 2014). The theory is rich in potential applications: for present purposes, most relevantly to language, communication, and reasoning (Misyak et al., 2014).

Both Game Theory and Virtual Bargaining allow us to explore cooperation and conflict in pragmatics and argumentation. Understanding how far cooperation is necessary, and conflict is possible, is an important goal in light of the discussion above. Traditional theory has assumed that pragmatics is fundamentally cooperative; but pragmatics can clearly happen in non-cooperative situations such as personal arguments. Conversely, we might assume that argumentation is fundamentally about conflict – about resolving differences of opinion, with one side winning. But argumentation theory increasingly stresses that some kind of cooperation is essential for resolving disputes. Tools such as Game Theory and Virtual Bargaining should allow us to develop theories of what we need to cooperate on and how far we can be in conflict while still successfully inferring meanings and resolving disputes.

A related body of research also treats pragmatics as rational social action, and has produced testable psychological accounts of pragmatics. This research uses Bayesian cognitive modelling, the main focus being a model known as the Rational Speech Act Model (henceforth, 'RSA' model) (e.g. Frank & Goodman, 2012;

Goodman & Frank, 2016; Goodman & Lassiter, in press; Goodman & Stuhlmüller, 2013). This research allows us to further reconsider the relationship between pragmatics and sources. The RSA model builds on the work of Grice in viewing talking as ‘a special case or variety of purposive, indeed rational, behaviour’ (Grice, 1975, p. 47). The model treats pragmatics as inference under uncertainty: uncertainty about speakers’ intentions and beliefs, the purpose of discourse, word meanings, and so on (Goodman & Frank, 2016). On this view, pragmatics is recursive inference, defined in the following way. Listeners (L) infer the state of the world, w , given a speaker’s (S) utterance, u , using Bayes’ rule:

$$P_L(w | u) \propto P_S(u | w)P(w)$$

Listeners assume that a speaker is more or less rational and has ‘chosen her utterances in proportion to the utility she expects to gain’ (Goodman & Frank, 2016):

$$P_S(u | w) \propto \exp(\alpha U(u;w))$$

α , here, is a parameter which captures how rational a speaker is. A speaker’s utility depends on providing help to the listener: it depends on how likely a literal listener is to infer the state of the world:

$$U(u;w) = \log P_{Lit}(w | u)$$

This reference to the literal speaker prevents an infinite recursion. Finally, the literal listener also performs Bayesian updating, assuming that the literal meaning of the utterance is true:

$$P_{Lit}(w | u) \propto \delta_{[u](w)} P(w)$$

The literal meaning is provided by a truth-functional semantics.

By formalizing pragmatic inference in this way, the RSA model can make specific predictions about experimental tasks. Experimental evidence provides increasing support for the model and its variants. The basic model fares well with

RATIONALITY, PRAGMATICS, AND SOURCES

referential communication tasks (Frank & Goodman, 2012). It can be supplemented with more elaborate utility functions which assign, for instance, disutility to production effort (Goodman & Frank, 2016); these more elaborate models can capture participants' sensitivity to production costs (Bergen, Goodman, & Levy, 2012). Forms of the model can also capture speakers' knowledge, and can accurately predict, in context, when people will derive scalar inferences (Goodman & Stuhlmüller, 2013). A parameter can also be added to the model to capture uncertainty about the speaker. The model thus allows joint inference, and can thereby accurately predict participants' behaviour with hyperbole; irony; simple metaphors (Goodman & Frank, 2016); vague predicates, such as 'tall' (Lassiter & Goodman, 2015); and embedded implicatures (Bergen, Levy, & Goodman, 2016). Thus far, the RSA approach applies at the level of individual utterances; a fuller account will need to consider wider discourse (Goodman & Frank, 2016). Plausibly, though, the RSA model could be used alongside the Bayes nets approach suggested above.

It is central to the RSA model that speakers are more or less rational, an assumption which is also implicit in Macagno and Walton's (2013) argumentative pragmatics. This assumption highlights the link between pragmatics and sources. There is reason to doubt that speakers are always rational, since they often depart from ideal behaviour at least when they are under time pressure and cognitive demands (Goodman & Frank, 2016). More work is needed to test this assumption and, as we have seen previously, to test what happens to pragmatic inference when speakers are perceived as irrational, unreliable or untrustworthy. But the RSA approach raises the prospect of a model which can integrate information about speakers. The model can currently integrate information about speakers' perceived

knowledge. It could profitably be extended to include information about speakers' past behaviours, personalities, and so on, much as Hilton (1995) suggested in his Attributional Model of pragmatic inference. If this can be done with a simple set of equations, then agent-based modelling could be used to explore the rationality of such behaviours on a population-wide level.

These new modelling methods – both game-theoretic and Bayesian – encourage a new specificity to pragmatic theories and potentially also to theories of argumentation. The methods also offer rich prospects for understanding the relationship between rationality, pragmatics, and sources. As these models are developed further and tested empirically, they should help to inform the large literature on the psychology of rationality, allowing us to interpret better participants' behaviours in experimental tasks.

8.3 Conclusions

This thesis has highlighted the complex interlocking of rationality, pragmatics, and sources. It has argued – through literature reviews, conceptual work, and experimentation – that these components are inseparable. The General Discussion has argued that pragmatics *is* rational behaviour: pragmatics draws on the same inference forms and broader argument forms as other rational behaviours, albeit to slightly different ends. And pragmatic theories are shot through with assumptions about sources. The triangular scheme is completed by the link between sources and rationality. People are sensitive to source information and can treat other people's claims sceptically. Questions remain about how rational this behaviour is, on a population level and across time. This thesis has suggested numerous avenues for future research. Recent developments permit a new unification of research on

RATIONALITY, PRAGMATICS, AND SOURCES

rationality, pragmatics, and sources. In such research, it will be important, when focusing on any one component, to keep the others in the back of one's mind.

Bibliography

- Adams, E. W. (1965). The logic of conditionals. *Inquiry*, 8(1–4), 166–197.
<https://doi.org/10.1080/00201746508601430>
- Adams, E. W. (1970). Subjunctive and Indicative Conditionals. *Foundations of Language*, 6(1), 89–94.
- Adams, E. W. (1975). *The Logic of Conditionals: An Application of Probability to Deductive Logic*. New York, NY: Springer Science & Business Media.
- Adler, J. (2015). Epistemological problems of testimony. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2015). Retrieved from
<https://plato.stanford.edu/archives/sum2015/entries/testimony-episprob/>
- Agnoli, F., & Krantz, D. H. (1989). Suppressing natural heuristics by formal instruction: The case of the conjunction fallacy. *Cognitive Psychology*, 21(4), 515–550. [https://doi.org/10.1016/0010-0285\(89\)90017-0](https://doi.org/10.1016/0010-0285(89)90017-0)
- Anderson, J. R. (1990). *The Adaptive Character of Thought*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Anderson, J. R. (1991a). Is human cognition adaptive? *Behavioral and Brain Sciences*, 14(3), 471–485. <https://doi.org/10.1017/S0140525X00070801>
- Anderson, J. R. (1991b). The adaptive nature of human categorization. *Psychological Review*, 98(3), 409–429. <https://doi.org/10.1037/0033-295X.98.3.409>
- Apanovitch, A. M., McCarthy, D., & Salovey, P. (2003). Using message framing to motivate HIV testing among low-income, ethnic minority women. *Health Psychology*, 22(1), 60–67. <https://doi.org/10.1037/0278-6133.22.1.60>

Arthur, R. T. W. (2011). *Natural Deduction: An Introduction to Logic with Real Arguments, a Little History, and Some Humour*. Peterborough, Ontario: Broadview.

Artz, D., & Gil, Y. (2007). A survey of trust in computer science and the Semantic Web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2), 58–71. <https://doi.org/10.1016/j.websem.2007.03.002>

Atanasov, P. D., Rescober, P., Stone, E., Swift, S. A., Servan-Schreiber, E., Tetlock, P. E., ... Mellers, B. (2017). Distilling the Wisdom of Crowds: Prediction Markets versus Prediction Polls. *Management Science*, 63(3), 691–706. <https://doi.org/10.1287/mnsc.2015.2374>

Aust, F., Diederhofen, B., Ullrich, S., & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods*, 45(2), 527–535. <https://doi.org/10.3758/s13428-012-0265-2>

Austin, J. L. (1965). *How to do things with words: the William James lectures delivered at Harvard Univ. in 1955*. Oxford, England: Clarendon Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>

Baratgin, J., Douven, I., Evans, J. S. B. T., Oaksford, M., Over, D., & Politzer, G. (2015). The New Paradigm and Mental Models. *Trends in Cognitive Sciences*, 19(10), 547–548. <https://doi.org/10.1016/j.tics.2015.06.013>

Bar-Hillel, M., & Neter, E. (1993). How alike is it versus how likely is it: A disjunction fallacy in probability judgments. *Journal of Personality and Social Psychology*, 65(6), 1119–1131. <https://doi.org/10.1037/0022-3514.65.6.1119>

RATIONALITY, PRAGMATICS, AND SOURCES

- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5–17.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Becker, G. M., Degroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science, 9*(3), 226–232. <https://doi.org/10.1002/bs.3830090304>
- Bennett, J. F. (2003). *A Philosophical Guide to Conditionals*. Oxford, England: Clarendon Press.
- Bergen, L., Goodman, N. D., & Levy, R. (2012). That’s what she (could have) said: How alternative utterances affect language use. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Bergen, L., Levy, R., & Goodman, N. D. (2016). Pragmatic reasoning through semantic inference. *Semantics and Pragmatics, 9*. <https://doi.org/10.3765/sp.9.20>
- Birnbaum, M. H., & Mellers, B. A. (1983). Bayesian inference: Combining base rates with opinions of sources who vary in credibility. *Journal of Personality*

and Social Psychology, 45(4), 792–804. <https://doi.org/10.1037/0022-3514.45.4.792>

Birnbaum, M. H., & Stegner, S. E. (1979). Source credibility in social judgment: Bias, expertise, and the judge's point of view. *Journal of Personality and Social Psychology*, 37(1), 48–74. <https://doi.org/10.1037/0022-3514.37.1.48>

Birnbaum, M. H., Wong, R., & Wong, L. K. (1976). Combining information from sources that vary in credibility. *Memory & Cognition*, 4(3), 330–336. <https://doi.org/10.3758/BF03213185>

Block, L. G., & Keller, P. A. (1995). When to Accentuate the Negative: The Effects of Perceived Efficacy and Message Framing on Intentions to Perform a Health-Related Behavior. *Journal of Marketing Research*, 32(2), 192. <https://doi.org/10.2307/3152047>

Bonini, N., Tentori, K., & Osherson, D. (2004). A Different Conjunction Fallacy. *Mind and Language*, 19(2), 199–210. <https://doi.org/10.1111/j.1468-0017.2004.00254.x>

Bonnefon, J.-F. (2009). A theory of utility conditionals: Paralogical reasoning from decision-theoretic leakage. *Psychological Review*, 116(4), 888–907. <https://doi.org/10.1037/a0017186>

Bonnefon, J.-F. (2012). Utility conditionals as consequential arguments: A random sampling experiment. *Thinking & Reasoning*, 18(3), 379–393. <https://doi.org/10.1080/13546783.2012.670751>

Bonnefon, J.-F. (2013). New ambitions for a new paradigm: Putting the psychology of reasoning at the service of humanity. *Thinking & Reasoning*, 19(3–4), 381–398. <https://doi.org/10.1080/13546783.2013.774294>

RATIONALITY, PRAGMATICS, AND SOURCES

- Bonnefon, J.-F. (2016). Predicting behavior on the basis of arguments from consequences. In F. Paglieri, L. Bonelli, & S. Felletti (Eds.), *The psychology of argument: Cognitive approaches to argumentation and persuasion*. (Vol. 59, pp. 1–16). London: College Publications.
- Bonnefon, J.-F., Girotto, V., & Legrenzi, P. (2011). The psychology of reasoning about preferences and unsequential decisions. *Synthese*, 185(1), 27–41. <https://doi.org/10.1007/s11229-011-9957-x>
- Bonnefon, J.-F., Haigh, M., & Stewart, A. J. (2013). Utility templates for the interpretation of conditional statements. *Journal of Memory and Language*, 68(4), 350–361. <https://doi.org/10.1016/j.jml.2013.01.002>
- Bonnefon, J.-F., & Hilton, D. J. (2004). Consequential Conditionals: Invited and Suppressed Inferences From Valued Outcomes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1), 28–37. <https://doi.org/10.1037/0278-7393.30.1.28>
- Bonnefon, J.-F., & Politzer, G. (2010). Pragmatic conditionals, conditional pragmatics, and the pragmatic component of reasoning. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking*. Oxford, England: Oxford University Press.
- Bonnefon, J.-F., & Sloman, S. A. (2013). The Causal Structure of Utility Conditionals. *Cognitive Science*, 37(1), 193–209. <https://doi.org/10.1111/cogs.12014>
- Bonnefon, J.-F., & Villejoubert, G. (2006). Tactful or Doubtful? Expectations of Politeness Explain the Severity Bias in the Interpretation of Probability Phrases. *Psychological Science*, 17(9), 747–751. <https://doi.org/10.1111/j.1467-9280.2006.01776.x>

RATIONALITY, PRAGMATICS, AND SOURCES

- Bonnefond, M., Van der Henst, J.-B., Gougain, M., Robic, S., Olsen, M. D., Weiss, O., & Noveck, I. A. (2012). How pragmatic interpretations arise from conditionals: Profiling the Affirmation of the Consequent argument with reaction time and EEG measures. *Journal of Memory and Language*, *67*(4), 468–485. <https://doi.org/10.1016/j.jml.2012.07.007>
- Bovens, L., & Hartmann, S. (2003). *Bayesian Epistemology* (OUP Catalogue). Oxford, England: Oxford University Press. Retrieved from <https://ideas.repec.org/b/oxp/obooks/9780199270408.html>
- Braine, M. D. S., & O'Brien, D. P. (1991). A theory of if: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, *98*(2), 182–203. <https://doi.org/10.1037/0033-295X.98.2.182>
- Braine, M. D. S., & O'Brien, D. P. (1998). *Mental Logic*. Hove, England: Psychology Press.
- Breheny, R. (2006). Communication and Folk Psychology. *Mind & Language*, *21*(1), 74–107. <https://doi.org/10.1111/j.1468-0017.2006.00307.x>
- Breheny, R., Ferguson, H. J., & Katsos, N. (2013). Investigating the timecourse of accessing conversational implicatures during incremental sentence interpretation. *Language and Cognitive Processes*, *28*(4), 443–467. <https://doi.org/10.1080/01690965.2011.649040>
- Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition*, *100*(3), 434–463. <https://doi.org/10.1016/j.cognition.2005.07.003>
- Brendl, C. M., Tory, E., & Lemm, K. M. (1995). Sensitivity to varying gains and losses: The role of self- discrepancies and event framing. *Journal of*

RATIONALITY, PRAGMATICS, AND SOURCES

Personality and Social Psychology, 69(6), 1028–1051.

<https://doi.org/10.1037/0022-3514.69.6.1028>

Briñol, P., & Petty, R. E. (2009). Source factors in persuasion: A self-validation approach. *European Review of Social Psychology*, 20(1), 49–96.

<https://doi.org/10.1080/10463280802643640>

Briñol, P., Petty, R. E., & Tormala, Z. L. (2004). Self-Validation of Cognitive Responses to Advertisements. *Journal of Consumer Research*, 30(4), 559–573. <https://doi.org/10.1086/380289>

Brown, P., & Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge, England: Cambridge University Press.

Busemeyer, J. R., & Bruza, P. D. (2012). *Quantum Models of Cognition and Decision*. Cambridge, England: Cambridge University Press.

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

<https://doi.org/10.1037/0022-3514.42.1.116>

Cann, R. (1993). *Formal Semantics: An Introduction*. Cambridge, England: Cambridge University Press.

Carston, R. (2008). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Hoboken, N.J.: John Wiley & Sons.

Cesario, J., Grant, H., & Tory, E. (2004). Regulatory Fit and Persuasion: Transfer From “Feeling Right.” *Journal of Personality and Social Psychology*, 86(3),

388–404. <https://doi.org/10.1037/0022-3514.86.3.388>

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and*

RATIONALITY, PRAGMATICS, AND SOURCES

Social Psychology, 39(5), 752–766. <https://doi.org/10.1037/0022-3514.39.5.752>

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology*, 66(3), 460–473. <https://doi.org/10.1037/0022-3514.66.3.460>

Chance, Z., Norton, M. I., Gino, F., & Ariely, D. (2011). Temporal view of the costs and benefits of self-deception. *Proceedings of the National Academy of Sciences*, 108(Supplement 3), 15655–15659. <https://doi.org/10.1073/pnas.1010658108>

Chick, C. F., Reyna, V. F., & Corbin, J. C. (2016). Framing effects are robust to linguistic disambiguation: A critical test of contemporary theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(2), 238–256. <https://doi.org/10.1037/xlm0000158>

Christmann, U., Mischo, C., & Flender, J. (2000). Argumentational Integrity: A Training Program for Dealing with Unfair Argumentative Contributions. *Argumentation*, 14(4), 339–360. <https://doi.org/10.1023/A:1007889322192>

Christmann, U., Mischo, C., & Groeben, N. (2000). Components of the Evaluation of Integrity Violations in Argumentative Discussions Relevant Factors and their Relationships. *Journal of Language and Social Psychology*, 19(3), 315–341. <https://doi.org/10.1177/0261927X00019003003>

Clark, B. (2013). *Relevance Theory*. Cambridge, England: Cambridge University Press.

RATIONALITY, PRAGMATICS, AND SOURCES

- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12(4), 335–359. [https://doi.org/10.1016/S0022-5371\(73\)80014-3](https://doi.org/10.1016/S0022-5371(73)80014-3)
- Clark, H. H. (1974). Semantics and comprehension. In T. A. Sebeok (Ed.) (Vol. 12, pp. 1291–1428). The Hague, The Netherlands: Mouton.
- Clark, H. H. (1996). *Using Language*. Cambridge, England: Cambridge University Press.
- Clark, R. (2011). *Meaningful Games: Exploring Language with Game Theory*. Cambridge, MA: MIT Press.
- Clifton, C., & Frazier, L. (2012). Discourse Integration Guided by the “Question under Discussion.” *Cognitive Psychology*, 65(2), 352–379. <https://doi.org/10.1016/j.cogpsych.2012.04.001>
- Coady, C. A. J. (1992). *Testimony: A philosophical study*. Oxford, England: Oxford University Press.
- Collins, P. J., & Hahn, U. (2016). Arguments and their sources. In F. Paglieri, L. Bonelli, & S. Felletti (Eds.), *The Psychology of Argument: Cognitive Approaches to Argumentation and Persuasion* (Vol. 59, pp. 129–149). London, England: College Publications.
- Collins, P. J., Hahn, U., von Gerber, Y., & Olsson, E. J. (2015). The bi-directional relationship between source characteristics and message content. In D. C. Noelle, R. Dale, S. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 423–428). Austin, TX: Cognitive Science Society.

RATIONALITY, PRAGMATICS, AND SOURCES

- Collins, P. J., Hahn, U., von Gerber, Y., & Olsson, E. J. (2017). *The bidirectional relationship between message content and source characteristics*. Manuscript submitted for publication.
- Collins, P. J., Krzyżanowska, K., Wheeler, G., Hartmann, S., & Hahn, U. (2017). *Conditionals and Testimony*. In Preparation.
- Condorcet, N. C. d. (1785). *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*. Paris, France.
- Cook, J., & Lewandowsky, S. (2011). *The Debunking Handbook*. St Lucia, Australia: University of Queensland.
- Corner, A., & Hahn, U. (2010). Message Framing, Normative Advocacy and Persuasive Success. *Argumentation*, 24(2), 153–163.
<https://doi.org/10.1007/s10503-009-9159-6>
- Corner, A., Hahn, U., & Oaksford, M. (2011). The psychological mechanism of the slippery slope argument. *Journal of Memory and Language*, 64(2), 133–152.
<https://doi.org/10.1016/j.jml.2010.10.002>
- Costa, A., Foucart, A., Arnon, I., Aparici, M., & Apesteguia, J. (2014). “Piensa” twice: On the foreign language effect in decision making. *Cognition*, 130(2), 236–254. <https://doi.org/10.1016/j.cognition.2013.11.010>
- Crupi, V., & Girotto, V. (2014). From is to ought, and back: how normative concerns foster progress in reasoning research. *Frontiers in Psychology*, 5.
<https://doi.org/10.3389/fpsyg.2014.00219>
- Crupi, V., Tentori, K., & Lombardi, L. (2009). Pseudodiagnosticity revisited. *Psychological Review*, 116(4), 971–985. <https://doi.org/10.1037/a0017050>

RATIONALITY, PRAGMATICS, AND SOURCES

Cruz, N., Baratgin, J., Oaksford, M., & Over, D. E. (2015). Bayesian reasoning with ifs and ands and ors. *Frontiers in Psychology, 6*.

<https://doi.org/10.3389/fpsyg.2015.00192>

Cruz, N., Over, D. E., Oaksford, M., & Baratgin, J. (2016). Centering and the meaning of conditionals. In A. Papafragou, D. J. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th annual conference of the cognitive science society*. Austin, TX: Cognitive Science Society.

Cumming, G. (2012). *Understanding The New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. London, England: Routledge.

Cummings, L. (2017). *Research in Clinical Pragmatics*. New York, NY: Springer.

Davidson, D. (1984). Radical interpretation. In *Inquiries into truth and interpretation* (pp. 125–140). Oxford, England: Clarendon Press.

Declerck, R., & Reed, S. (2001). *Conditionals: A Comprehensive Empirical Analysis*. Berlin: Walter de Gruyter.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal Versus Naive Diversification: How Inefficient is the $1/N$ Portfolio Strategy? *Review of Financial Studies, 22*(5), 1915–1953. <https://doi.org/10.1093/rfs/hhm075>

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118. <https://doi.org/10.1037/0033-2909.129.1.74>

Detweiler, J. B., Bedell, B. T., Salovey, P., Pronin, E., & Rothman, A. J. (1999). Message framing and sunscreen use: Gain-framed messages motivate beachgoers. *Health Psychology, 18*(2), 189–196. <https://doi.org/10.1037/0278-6133.18.2.189>

RATIONALITY, PRAGMATICS, AND SOURCES

- Dodd, D. H., & Bradshaw, J. M. (1980). Leading questions and memory: Pragmatic constraints. *Journal of Verbal Learning and Verbal Behavior*, *19*(6), 695–704. [https://doi.org/10.1016/S0022-5371\(80\)90379-5](https://doi.org/10.1016/S0022-5371(80)90379-5)
- Douven, I. (2010). The pragmatics of belief. *Journal of Pragmatics*, *42*(1), 35–47. <https://doi.org/10.1016/j.pragma.2009.05.025>
- Douven, I. (2011). A role for normativism. *Behavioral and Brain Sciences*, *34*(5), 252–253. <https://doi.org/10.1017/S0140525X11000471>
- Douven, I. (2012). Learning Conditional Information. *Mind & Language*, *27*(3), 239–263. <https://doi.org/10.1111/j.1468-0017.2012.01443.x>
- Douven, I. (2015). *The Epistemology of Indicative Conditionals: Formal and Empirical Approaches*. Cambridge, England: Cambridge University Press.
- Douven, I., & Krzyżanowska, K. (2017). *The semantics-pragmatics interface: an empirical investigation*. Manuscript submitted for publication.
- Douven, I., & Verbrugge, S. (2010). The Adams family. *Cognition*, *117*(3), 302–318. <https://doi.org/10.1016/j.cognition.2010.08.015>
- Druckman, J. N., & McDermott, R. (2008). Emotion and the Framing of Risky Choice. *Political Behavior*, *30*(3), 297–321. <https://doi.org/10.1007/s11109-008-9056-y>
- Dulany, D. E., & Hilton, D. J. (1991). Conversational Implicature, Conscious Representation, and the Conjunction Fallacy. *Social Cognition*, *9*(1), 85–110. <https://doi.org/10.1521/soco.1991.9.1.85>
- Edgington, D. (1995). On Conditionals. *Mind*, *104*(414), 235–329. <https://doi.org/10.1093/mind/104.414.235>

RATIONALITY, PRAGMATICS, AND SOURCES

- Edgington, D. (2014). Indicative conditionals. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition). Retrieved from <https://plato.stanford.edu/archives/win2014/entries/conditionals/>
- Eemeren, F. H. van, & Garssen, B. (2013). Viewing the study of argumentation as normative pragmatics. In A. Capone, F. L. Piparo, & M. Carapezza (Eds.), *Perspectives on Pragmatics and Philosophy* (pp. 515–536). New York, NY: Springer.
- Eemeren, F. H. van, Garssen, B., & Meuffels, B. (2009). *Fallacies and Judgments of Reasonableness: Empirical Research Concerning the Pragma-Dialectical Discussion Rules* (Vol. 16). Dordrecht: Springer.
- Eemeren, F. H. van, Garssen, B., & Meuffels, B. (2012). The disguised abusive ad hominem empirically investigated: Strategic manoeuvring with direct personal attacks. *Thinking & Reasoning*, 18(3), 344–364.
<https://doi.org/10.1080/13546783.2012.678666>
- Eemeren, F. H. van, & Grootendorst, R. (1992). *Argumentation, communication, and fallacies: A pragma-dialectical perspective*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Eemeren, F. H. van, & Grootendorst, R. (1995). The Pragma-Dialectical Approach to Fallacies. In H. V. Hansen & R. C. Pinto (Eds.), *Fallacies: Classical and Contemporary Readings*. Philadelphia: Pennsylvania State University Press.
- Eemeren, F. H. van, & Grootendorst, R. (2004). *A systematic theory of argumentation: The pragma-dialectical approach*. Cambridge, England: Cambridge University Press.

RATIONALITY, PRAGMATICS, AND SOURCES

- Eemeren, F. H. van, & Houtlosser, P. (1999). Strategic Manoeuvring in Argumentative Discourse. *Discourse Studies*, 1(4), 479–497.
<https://doi.org/10.1177/1461445699001004005>
- Elder, C.-H., & Jaszczolt, K. M. (2016). Towards a pragmatic category of conditionals. *Journal of Pragmatics*, 98, 36–53.
<https://doi.org/10.1016/j.pragma.2016.04.013>
- Elqayam, S., & Evans, J. S. B. T. (2011). Subtracting “ought” from “is”: descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences*, 34, 233–248. <https://doi.org/doi:10.1080/13546783.2013.834268>
- Evans, J. S. B. T., Handley, S. J., Neilens, H., & Over, D. E. (2007). Thinking about conditionals: A study of individual differences. *Memory & Cognition*, 35(7), 1772–1784. <https://doi.org/10.3758/BF03193509>
- Evans, J. S. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(2), 321–335. <https://doi.org/10.1037/0278-7393.29.2.321>
- Evans, J. S. B. T., Neilens, H., Handley, S. J., & Over, D. E. (2008). When can we say “if”? *Cognition*, 108(1), 100–116.
<https://doi.org/10.1016/j.cognition.2008.02.001>
- Evans, J. S. B. T., & Over, D. E. (2004). *If: Supposition, Pragmatics, and Dual Processes*. Oxford, England: Oxford University Press.
- Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, 8(3), 223–241. <https://doi.org/10.1177/1745691612460685>

RATIONALITY, PRAGMATICS, AND SOURCES

- Fenton, N., Neil, M., & Lagnado, D. A. (2013). A General Structure for Legal Arguments About Evidence Using Bayesian Networks. *Cognitive Science*, 37(1), 61–102. <https://doi.org/10.1111/cogs.12004>
- Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research*, 50(2), 123–129. <https://doi.org/10.1007/BF00309212>
- Finch, W. H., Bolin, J. E., & Kelley, K. (2016). *Multilevel Modeling Using R*. CRC Press.
- Frank, M. C., & Goodman, N. D. (2012). Predicting Pragmatic Reasoning in Language Games. *Science*, 336(6084), 998–998. <https://doi.org/10.1126/science.1218633>
- Franke, M., & Jäger, G. (2016). Probabilistic pragmatics, or why Bayes' rule is probably important for pragmatics. *Zeitschrift Für Sprachwissenschaft*, 35(1). <https://doi.org/10.1515/zfs-2016-0002>
- Friedman, R. D. (1987). Route Analysis of Credibility and Hearsay. *The Yale Law Journal*, 96(4), 667–742. <https://doi.org/10.2307/796360>
- Fugard, A. J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2011). How people interpret conditionals: Shifts toward the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(3), 635–648. <https://doi.org/10.1037/a0022329>
- Gallagher, K. M., & Updegraff, J. A. (2012). Health message framing effects on attitudes, intentions, and behavior: a meta-analytic review. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, 43(1), 101–116. <https://doi.org/10.1007/s12160-011-9308-7>

RATIONALITY, PRAGMATICS, AND SOURCES

Galton, F. (1907). Vox Populi. *Nature*, 75, 450–451.

<https://doi.org/10.1038/075450a0>

Ganea, P. A., Koenig, M. A., & Millett, K. G. (2011). Changing your mind about things unseen: Toddlers' sensitivity to prior reliability. *Journal of Experimental Child Psychology*, 109(4), 445–453.

<https://doi.org/10.1016/j.jecp.2011.02.011>

Garber, D. (1983). Old Evidence and Logical Omniscience in Bayesian Confirmation Theory. In J. Earman (Ed.), *Testing Scientific Theories* (Vol. 10, pp. 99–131). Minneapolis, MN: University of Minnesota Press.

Geis, M. L., & Zwicky, A. M. (1971). On Invited Inferences. *Linguistic Inquiry*, 2(4), 561–566.

Gerend, M. A., & Cullen, M. (2008). Effects of message framing and temporal context on college student drinking behavior. *Journal of Experimental Social Psychology*, 44(4), 1167–1173. <https://doi.org/10.1016/j.jesp.2008.02.007>

Geurts, B. (2013). Alternatives in Framing and Decision Making. *Mind & Language*, 28(1), 1–19. <https://doi.org/10.1111/mila.12005>

Geurts, B., Katsos, N., Cummins, C., Moons, J., & Noordman, L. (2010). Scalar quantifiers: Logic, acquisition, and processing. *Language and Cognitive Processes*, 25(1), 130–148. <https://doi.org/10.1080/01690960902955010>

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

<https://doi.org/10.1037/0033-295X.103.3.592>

Gigerenzer, G. (2000). *Adaptive Thinking: Rationality in the Real World*. Oxford, England: Oxford University Press.

RATIONALITY, PRAGMATICS, AND SOURCES

Gigerenzer, G. (2010). *Rationality for Mortals: How People Cope with Uncertainty*.

Oxford, England: Oxford University Press.

Gigerenzer, G. (2015). On the Supposed Evidence for Libertarian Paternalism.

Review of Philosophy and Psychology, 1–23. <https://doi.org/10.1007/s13164-015-0248-1>

Gigerenzer, G., & Selten, R. (2002). *Bounded Rationality: The Adaptive Toolbox*.

Cambridge, MA: MIT Press.

Gigerenzer, G., Todd, P. M., & Gerd Gigerenzer, A. R. (1999). *Simple Heuristics That Make*

Us Smart. Oxford, England: Oxford University Press.

Gigone, D., & Hastie, R. (1997). Proper analysis of the accuracy of group judgments.

Psychological Bulletin, 121(1), 149–167. <https://doi.org/10.1037/0033-2909.121.1.149>

Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable:

Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, 59(4), 601–613. <https://doi.org/10.1037/0022-3514.59.4.601>

Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe

everything you read. *Journal of Personality and Social Psychology*, 65(2), 221–233. <https://doi.org/10.1037/0022-3514.65.2.221>

Gilbert, N. (2008). *Agent-Based Models*. London, England: Sage.

Giora, R. (1999). On the priority of salient meanings: Studies of literal and figurative

language. *Journal of Pragmatics*, 31(7), 919–929. [https://doi.org/10.1016/S0378-2166\(98\)00100-3](https://doi.org/10.1016/S0378-2166(98)00100-3)

Giora, R. (2002). Literal vs. figurative language: Different or equal? *Journal of*

Pragmatics, 34(4), 487–506. [https://doi.org/10.1016/S0378-2166\(01\)00045-5](https://doi.org/10.1016/S0378-2166(01)00045-5)

RATIONALITY, PRAGMATICS, AND SOURCES

- Giora, R. (2009). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics (Includes Cognitive Linguistic Bibliography)*, 8(3), 183–206. <https://doi.org/10.1515/cogl.1997.8.3.183>
- Giorotto, V., Kemmelmeier, M., Sperber, D., & van der Henst, J.-B. (2001). Inept reasoners or pragmatic virtuosos? Relevance and the deontic selection task. *Cognition*, 81(2), B69–B76. [https://doi.org/doi:10.1016/S0010-0277\(01\)00124-X](https://doi.org/doi:10.1016/S0010-0277(01)00124-X)
- Glass, D. H. (2007). Coherence measures and inference to the best explanation. *Synthese*, 157(3), 275–296. <https://doi.org/10.1007/s11229-006-9055-7>
- Glymour, C. N. (1980). *Theory and Evidence*. Princeton, NJ: Princeton University Press.
- Goodman, N. D., & Frank, M. C. (2016). Pragmatic Language Interpretation as Probabilistic Inference. *Trends in Cognitive Sciences*, 20(11), 818–829. <https://doi.org/10.1016/j.tics.2016.08.005>
- Goodman, N. D., & Lassiter, D. (in press). Probabilistic Semantics and Pragmatics: Uncertainty in Language and Thought. In S. Lappin & C. Fox (Eds.), *Handbook of Contemporary Semantic Theory* (2nd Edition). Hoboken, N.J.: Wiley-Blackwell.
- Goodman, N. D., & Stuhlmüller, A. (2013). Knowledge and Implicature: Modeling Language Understanding as Social Cognition. *Topics in Cognitive Science*, 5(1), 173–184. <https://doi.org/10.1111/tops.12007>
- Goodwin, G. P. (2014). Is the basic conditional probabilistic? *Journal of Experimental Psychology: General*, 143(3), 1214–1241. <https://doi.org/10.1037/a0034232>

RATIONALITY, PRAGMATICS, AND SOURCES

- Goodwin, J. (2001). The non-cooperative pragmatics of arguing. In E. Nemeth (Ed.), *Pragmatics in 2000: Selected papers from the 7th International Pragmatics Conference*. (Vol. 2, pp. 263–277). Antwerp, Belgium: International Pragmatics Association.
- Graham, G. (2014). *The Disordered Mind: An Introduction to Philosophy of Mind and Mental Illness*. London, England: Routledge.
- Grice, H. . (1975). Logic and Conversation. In Davidson, D. & (second) Harman, G. (Eds.), *The logic of grammar* (pp. 64–75). Encino, CA: Dickenson.
- Grice, H. P. (1957). Meaning. In *Philosophical Review*, Volume 66 (reprinted in Grice, 1989, 213-223).
- Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). “Some,” and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, *116*(1), 42–55.
<https://doi.org/10.1016/j.cognition.2010.03.014>
- Groome, D. (2013). *An Introduction to Cognitive Psychology: Processes and disorders*. Hove, England: Psychology Press.
- Gruenfeld, D. H., & Wyer, R. S. (1992). Semantics and pragmatics of social influence: How affirmations and denials affect beliefs in referent propositions. *Journal of Personality and Social Psychology*, *62*(1), 38–49.
<https://doi.org/10.1037/0022-3514.62.1.38>
- Hacking, I. (2001). *An Introduction to Probability and Inductive Logic*. Cambridge, England: Cambridge University Press.

RATIONALITY, PRAGMATICS, AND SOURCES

- Hahn, U. (2011). Why rational norms are indispensable. *Behavioral and Brain Sciences*, 34(5), 257–258. <https://doi.org/10.1017/S0140525X11000641>
- Hahn, U. (2014). The Bayesian boom: good thing or bad? *Cognitive Science*, 5, 765. <https://doi.org/10.3389/fpsyg.2014.00765>
- Hahn, U., & Harris, A. J. L. (2014). What does it mean to be biased: motivated reasoning and rationality. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 61, pp. 41–102). Cambridge, MA: Academic Press.
- Hahn, U., Harris, A. J. L., & Corner, A. (2016). Public Reception of Climate Science: Coherence, Reliability, and Independence. *Topics in Cognitive Science*, 8(1), 180–195. <https://doi.org/10.1111/tops.12173>
- Hahn, U., Harris, A. J. L., & Corner, A. J. (2009). Argument Content and Argument Source: An Exploration. *Informal Logic*, 29(4), 337–367.
- Hahn, U., Harris, A. J. L., & Oaksford, M. (2013). Rational argument, rational inference. *Argument & Computation*, 4(1), 21–35. <https://doi.org/10.1080/19462166.2012.689327>
- Hahn, U., & Hornikx, J. (2016). A normative framework for argument quality: argumentation schemes with a Bayesian foundation. *Synthese*, 193(6), 1833–1873. <https://doi.org/10.1007/s11229-015-0815-0>
- Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review*, 114(3), 704–732. <https://doi.org/10.1037/0033-295X.114.3.704>
- Hahn, U., Oaksford, M., & Harris, A. J. L. (2012). Testimony and Argument: A Bayesian Perspective. In F. Zenker (Ed.), *Bayesian Argumentation* (pp. 15–38). New York, NY: Springer.

RATIONALITY, PRAGMATICS, AND SOURCES

- Haigh, M., & Bonnefon, J.-F. (2015). Eye Movements Reveal How Readers Infer Intentions From the Beliefs and Desires of Others. *Experimental Psychology*, *62*(3), 206–213. <https://doi.org/10.1027/1618-3169/a000290>
- Haigh, M., Ferguson, H. J., & Stewart, A. J. (2014). An eye-tracking investigation into readers' sensitivity to actual versus expected utility in the comprehension of conditionals. *The Quarterly Journal of Experimental Psychology*, *67*(1), 166–185. <https://doi.org/10.1080/17470218.2013.797475>
- Haigh, M., & Stewart, A. J. (2011). The influence of clause order, congruency, and probability on the processing of conditionals. *Thinking & Reasoning*, *17*(4), 402–423. <https://doi.org/10.1080/10463283.2011.628000>
- Haigh, M., Stewart, A. J., & Connell, L. (2012). Reasoning as we read: Establishing the probability of causal conditionals. *Memory & Cognition*, *41*(1), 152–158. <https://doi.org/10.3758/s13421-012-0250-0>
- Harkins, S. G., & Petty, R. E. (1981). The Multiple Source Effect in Persuasion: The Effects of Distraction. *Personality and Social Psychology Bulletin*, *7*(4), 627–635. <https://doi.org/10.1177/014616728174019>
- Harley, T. A. (2013). *The Psychology of Language: From Data to Theory*. Hove, England: Psychology Press.
- Harris, A. J. L., Corner, A., & Hahn, U. (2013). James is polite and punctual (and useless): A Bayesian formalisation of faint praise. *Thinking & Reasoning*, *19*(3–4), 414–429. <https://doi.org/10.1080/13546783.2013.801367>
- Harris, A. J. L., & Hahn, U. (2009). Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(5), 1366–1373. <https://doi.org/10.1037/a0016567>

RATIONALITY, PRAGMATICS, AND SOURCES

- Harris, A. J. L., Hahn, U., Madsen, J. K., & Hsu, A. S. (2016). The Appeal to Expert Opinion: Quantitative Support for a Bayesian Network Approach. *Cognitive Science*, 40(6), 1496–1533. <https://doi.org/10.1111/cogs.12276>
- Hartmann, S., & Fitelson, B. (2015). A New Garber-Style Solution to the Problem of Old Evidence. *Philosophy of Science*, 82(4), 712–717. <https://doi.org/10.1086/682916>
- Hartmann, S., & Rafiee Rad, S. (2017). Learning indicative conditionals. *Manuscript Submitted for Publication*. Retrieved from <http://www.rafieread.org/research.html>
- Hasson, U., Simmons, J. P., & Todorov, A. (2005). Believe It or Not On the Possibility of Suspending Belief. *Psychological Science*, 16(7), 566–571. <https://doi.org/10.1111/j.0956-7976.2005.01576.x>
- Hertwig, R., Benz, B., & Krauss, S. (2008). The conjunction fallacy and the many meanings of and. *Cognition*, 108(3), 740–753. <https://doi.org/10.1016/j.cognition.2008.06.008>
- Hertwig, R., & Chase, V. M. (1998). Many Reasons or Just One: How Response Mode Affects Reasoning in the Conjunction Problem. *Thinking & Reasoning*, 4(4), 319–352. <https://doi.org/10.1080/135467898394102>
- Hertwig, R., & Gigerenzer, G. (1999). The “conjunction fallacy” revisited: how intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, 12(4), 275–305. [https://doi.org/10.1002/\(SICI\)1099-0771\(199912\)12:4<275::AID-BDM323>3.0.CO;2-M](https://doi.org/10.1002/(SICI)1099-0771(199912)12:4<275::AID-BDM323>3.0.CO;2-M)
- Heyman, G. D., & Legare, C. H. (2005). Children’s Evaluation of Sources of Information About Traits. *Developmental Psychology*, 41(4), 636–647. <https://doi.org/10.1037/0012-1649.41.4.636>

RATIONALITY, PRAGMATICS, AND SOURCES

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

<https://doi.org/10.1037/0033-2909.118.2.248>

Hilton, D. J., Kimmelmeier, M., & Bonnefon, J.-F. (2005). Putting Ifs to Work: Goal-Based Relevance in Conditional Directives. *Journal of Experimental Psychology: General*, *134*(3), 388–405. <https://doi.org/10.1037/0096-3445.134.3.388>

<https://doi.org/10.1037/0096-3445.134.3.388>

<https://doi.org/10.1037/0096-3445.134.3.388>

Holtgraves, T. M. (2002). *Language As Social Action: Social Psychology and Language Use*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Horn, L. R. (1984). Toward a new taxonomy for pragmatic inference. In D. Schiffrin (Ed.), *Meaning, form, and use in context: linguistic applications*. Washington, DC: Georgetown University Press.

Horn, L. R. (1989). *A natural history of negation*. Chicago, IL: University of Chicago Press.

Horn, L. R. (2004). Implicature. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 3–28). Hoboken, N.J.: Wiley-Blackwell.

Howson, C. (1991). The “Old Evidence” Problem. *The British Journal for the Philosophy of Science*, *42*(4), 547–555. <https://doi.org/10.1093/bjps/42.4.547>

Huang, Y. (2007). *Pragmatics*. Oxford, England: Oxford University Press.

Ingram, J., Hand, C. J., & Moxey, L. M. (2014). Processing inferences drawn from the logically equivalent frames half full and half empty. *Journal of Cognitive Psychology*, *26*(7), 799–817. <https://doi.org/10.1080/20445911.2014.956747>

Jackson, F. (1979). On Assertion and Indicative Conditionals. *The Philosophical Review*, *88*(4), 565. <https://doi.org/10.2307/2184845>

RATIONALITY, PRAGMATICS, AND SOURCES

- Jarvstad, A., & Hahn, U. (2011). Source Reliability and the Conjunction Fallacy. *Cognitive Science*, 35(4), 682–711. <https://doi.org/10.1111/j.1551-6709.2011.01170.x>
- Jaszczolt, K. (2007). *Default Semantics: Foundations of a Compositional Theory of Acts of Communication*. Oxford, England: Oxford University Press.
- Jern, A., Chang, K. K., & Kemp, C. (2014). Belief polarization is not always irrational. *Psychological Review*, 121(2), 206–224. <https://doi.org/10.1037/a0035941>
- Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Mahwah, New Jersey: Erlbaum.
- Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, 109(4), 646–678. <https://doi.org/10.1037/0033-295X.109.4.646>
- Johnson-Laird, P. N., Khemlani, S. S., & Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences*, 19(4), 201–214. <https://doi.org/10.1016/j.tics.2015.02.006>
- Kadane, J. B., & Schum, D. A. (1996). *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*. Mahwah, New Jersey: John Wiley & Sons.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. London, England: Penguin.
- Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, 64(6), 515–526. <https://doi.org/10.1037/a0016755>
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80(4), 237–251. <https://doi.org/10.1037/h0034747>

RATIONALITY, PRAGMATICS, AND SOURCES

- Katsikopoulos, K. V. (2014). Bounded rationality: the two cultures. *Journal of Economic Methodology*, 21(4), 361–374.
<https://doi.org/10.1080/1350178X.2014.965908>
- Kelman, H. C. (1958). Compliance, identification, and internalization: Three processes of attitude change. *Journal of Conflict Resolution*, 2(1). Retrieved from <http://scholar.harvard.edu/hckelman/publications/compliance-identification-and-internalization-three-processes-attitude-change>
- Kempson, R., Meyer-Viol, W., & Gabbay, D. M. (2000). *Dynamic Syntax: The Flow of Language Understanding*. Hoboken, N.J.: Wiley.
- Keren, G. (2007). Framing, intentions, and trust–choice incompatibility. *Organizational Behavior and Human Decision Processes*, 103(2), 238–255.
<https://doi.org/10.1016/j.obhdp.2007.02.002>
- Kerr, N. S., & Tindale, R. S. (2004). Group Performance and Decision Making. *Annual Review of Psychology*, 55(1), 623–655.
<https://doi.org/10.1146/annurev.psych.55.090902.142009>
- Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The Foreign-Language Effect Thinking in a Foreign Tongue Reduces Decision Biases. *Psychological Science*, 23(6), 661–668. <https://doi.org/10.1177/0956797611432178>
- Kolmogorov, A. N. (2013). *Foundations of the Theory of Probability*. Eastford, CT: Martino Fine Books.
- Korb, K. B., & Nicholson, A. E. (2011). *Bayesian Artificial Intelligence* (Second Edition). Boca Raton, Florida: CRC Press.
- Korta, K., & Perry, J. (2015). Pragmatics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2015 Edition). Retrieved from <https://plato.stanford.edu/archives/win2015/entries/pragmatics/>

RATIONALITY, PRAGMATICS, AND SOURCES

- Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, *142*(2), 573–603.
<https://doi.org/10.1037/a0029146>
- Kruschke, J. K. (2015). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Cambridge, MA: Academic Press.
- Krzyżanowska, K. (2015). *Between “If” and “Then”. Towards an empirically informed philosophy of conditionals*. (PhD). University of Groningen, Groningen, Netherlands. Retrieved from <http://karolinakrzyzanowska.com/>
- Krzyżanowska, K., Collins, P. J., & Hahn, U. (Accepted). *The puzzle of conditionals with true clauses: against the Gricean account*. To appear in Proceedings of the 39th Annual Meeting of the Cognitive Science Society, Austin, TX.
- Krzyżanowska, K., Collins, P. J., & Hahn, U. (2017). Between a conditional’s antecedent and its consequent: Discourse coherence vs. probabilistic relevance. *Cognition*, *164*, 199–205.
<https://doi.org/10.1016/j.cognition.2017.03.009>
- Kühberger, A. (1995). The Framing of Decisions: A New Look at Old Problems. *Organizational Behavior and Human Decision Processes*, *62*(2), 230–240.
<https://doi.org/10.1006/obhd.1995.1046>
- Kühberger, A., & Tanner, C. (2010). Risky choice framing: Task versions and a comparison of prospect theory and fuzzy-trace theory. *Journal of Behavioral Decision Making*, *23*(3), 314–329. <https://doi.org/10.1002/bdm.656>
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. <https://doi.org/10.1037/0033-2909.108.3.480>
- Lalor, K. M., & Hailey, B. J. (1989). The Effects of Message Framing and Feelings of Susceptibility to Breast Cancer on Reported Frequency of Breast Self-

- Examination. *International Quarterly of Community Health Education*, 10(3), 183–192. <https://doi.org/10.2190/GMFB-WYND-QJYA-8LJC>
- Lassiter, D., & Goodman, N. D. (2015). Adjectival vagueness in a Bayesian model of interpretation. *Synthese*, 1–36. <https://doi.org/10.1007/s11229-015-0786-1>
- Leahy, B. (2011). Presuppositions and Antipresuppositions in Conditionals. *Semantics and Linguistic Theory*, 21, 257. <https://doi.org/10.3765/salt.v21i0.2613>
- Lee, C. J. (2006). Gricean Charity: The Gricean Turn in Psychology. *Philosophy of the Social Sciences*, 36(2), 193–218. <https://doi.org/10.1177/0048393106287235>
- Lee, J. J., & Pinker, S. (2010). Rationales for indirect speech: The theory of the strategic speaker. *Psychological Review*, 117(3), 785–807. <https://doi.org/10.1037/a0019688>
- Leitgeb, H., & Pettigrew, R. (2010b). An objective justification of bayesianism ii: the consequences of minimizing inaccuracy*. *Philosophy of Science*, 77, 236–272. <https://doi.org/doi:10.1086/651318>
- Leitgeb, H., & Pettigrew, R. (2010). An Objective Justification of Bayesianism I: Measuring Inaccuracy. *Philosophy of Science*, 77(2), 201–235. <https://doi.org/10.1086/651317>
- Leith, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, 69(1), 1–33. <https://doi.org/10.18637/jss.v069.i01>
- Levin, I. P., Gaeth, G. ., Schreiber, J., & Lauriola, M. (2002). A New Look at Framing Effects: Distribution of Effect Sizes, Individual Differences, and Independence of Types of Effects. *Organizational Behavior and Human Decision Processes*, 88(1), 411–429. <https://doi.org/10.1006/obhd.2001.2983>

RATIONALITY, PRAGMATICS, AND SOURCES

- Levin, I. P., Schneider, S. L., & Gaeth, G. . (1998). All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects. *Organizational Behavior and Human Decision Processes*, 76(2), 149–188.
- Levine, G., & Parkinson, S. (1994). *Experimental Methods in Psychology*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge, England: Cambridge University Press.
- Levinson, S. C. (2000). *Presumptive meanings: the theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- List, C. (2004). On the Significance of the Absolute Margin. *The British Journal for the Philosophy of Science*, 55(3), 521–544.
<https://doi.org/10.1093/bjps/55.3.521>
- Littlepage, G. E., & Mueller, A. L. (1997). Recognition and utilization of expertise in problem-solving groups: Expert characteristics and behavior. *Group Dynamics: Theory, Research, and Practice*, 1(4), 324–328.
<https://doi.org/10.1037/1089-2699.1.4.324>
- Littlepage, G. E., Schmidt, G. W., Whisler, E. W., & Frost, A. G. (1995). An input-process-output analysis of influence and performance in problem-solving groups. *Journal of Personality and Social Psychology*, 69(5), 877–889.
<https://doi.org/10.1037/0022-3514.69.5.877>
- Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7(4), 560–572. [https://doi.org/10.1016/0010-0285\(75\)90023-7](https://doi.org/10.1016/0010-0285(75)90023-7)
- Macagno, F., & Walton, D. (2013). Implicatures as Forms of Argument. In A. Capone, F. L. Piparo, & M. Carapezza (Eds.), *Perspectives on Pragmatics*

and Philosophy (pp. 203–225). Springer International Publishing.

https://doi.org/10.1007/978-3-319-01011-3_9

Macdonald, R. R., & Gilhooly, K. J. (1990). More about *linda or* conjunctions in context. *European Journal of Cognitive Psychology*, 2(1), 57–70.

<https://doi.org/10.1080/09541449008406197>

Maio, G. R., & Esses, V. M. (2001). The need for affect: individual differences in the motivation to approach or avoid emotions. *Journal of Personality*, 69(4), 583–615.

Mair, P., Schoenbrodt, F., & Wilcox, R. R. (2017). *WRS2: Wilcox robust estimation and testing*.

Mandel, D. R. (2001). Gain-Loss Framing and Choice: Separating Outcome Formulations from Descriptor Formulations. *Organizational Behavior and Human Decision Processes*, 85(1), 56–76.

<https://doi.org/10.1006/obhd.2000.2932>

Mandel, D. R. (2014). Do framing effects reveal irrational choice? *Journal of Experimental Psychology: General*, 143(3), 1185–1198.

<https://doi.org/10.1037/a0034207>

Mandelbaum, E. (2014). Thinking is Believing. *Inquiry*, 57(1), 55–96.

<https://doi.org/10.1080/0020174X.2014.858417>

Manktelow, K. (2012). *Thinking and Reasoning: An Introduction to the Psychology of Reason, Judgment and Decision Making*. Hove, England: Psychology Press.

Mann, T. E., & Ornstein, N. J. (2013). *It's Even Worse Than It Looks: How the American Constitutional System Collided With the New Politics of Extremism*. New York, NY: Basic Books.

RATIONALITY, PRAGMATICS, AND SOURCES

- Marcu, D. (2000). Perlocutions: The Achilles' heel of speech act theory. *Journal of Pragmatics*, 32(12), 1719–1741. [https://doi.org/10.1016/S0378-2166\(99\)00121-6](https://doi.org/10.1016/S0378-2166(99)00121-6)
- Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, 112(3), 367–380. <https://doi.org/10.1016/j.cognition.2009.05.012>
- Matsui, T., & Fitneva, S. A. (2009). Knowing how we know: Evidentiality and cognitive development. *New Directions for Child and Adolescent Development*, 2009(125), 1–11. <https://doi.org/10.1002/cd.246>
- Mazar, N., Amir, O., & Ariely, D. (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research*, 45(6), 633–644. <https://doi.org/10.1509/jmkr.45.6.633>
- McCready, E. (2014). *Reliability in Pragmatics*. Oxford, England: Oxford University Press.
- McKenzie, C. R. M., & Nelson, J. D. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulletin & Review*, 10(3), 596–602. <https://doi.org/10.3758/BF03196520>
- Mellers, B., Stone, E., Atanasov, P., Rohrbaugh, N., Emlen, S., Ungar, L., ... Tetlock, P. (2015). The psychology of intelligence analysis: Drivers of prediction accuracy in world politics. *Journal of Experimental Psychology: Applied*, 21(1), 1–14. <https://doi.org/10.1037/xap0000040>
- Mellers, B., Stone, E., Murray, T., Minster, A., Rohrbaugh, N., Bishop, M., ... Tetlock, P. (2015). Identifying and Cultivating Superforecasters as a Method of Improving Probabilistic Predictions. *Perspectives on Psychological Science*, 10(3), 267–281. <https://doi.org/10.1177/1745691615577794>

RATIONALITY, PRAGMATICS, AND SOURCES

- Mellers, B., Ungar, L., Baron, J., Ramos, J., Gurcay, B., Fincher, K., ... Tetlock, P. E. (2014). Psychological Strategies for Winning a Geopolitical Forecasting Tournament. *Psychological Science*, *25*(5), 1106–1115.
<https://doi.org/10.1177/0956797614524255>
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*(2), 57–74.
<https://doi.org/10.1017/S0140525X10000968>
- Mercier, H., & Sperber, D. (2017). *The Enigma of Reason*. Cambridge, MA: Harvard University Press.
- Meredith, M., & Kruschke, J. K. (2013). BEST: Bayesian Estimation Supersedes the t-Test (Version .2.0). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://cran.r-project.org/web/packages/BEST>
- Messer, W. S., & Griggs, R. A. (1993). Another look at Linda. *Bulletin of the Psychonomic Society*, *31*(3), 193–196. <https://doi.org/10.3758/BF03337322>
- Mey, J. L. (2001). *Pragmatics: An Introduction*. Hoboken, N.J.: Wiley.
- Meyerowitz, B. E., & Chaiken, S. (1987). The effect of message framing on breast self-examination attitudes, intentions, and behavior. *Journal of Personality and Social Psychology*, *52*(3), 500–510. <https://doi.org/10.1037/0022-3514.52.3.500>
- Meyers-Levy, J., & Maheswaran, D. (2004). Exploring Message Framing Outcomes When Systematic, Heuristic, or Both Types of Processing Occur. *Journal of Consumer Psychology*, *14*(1–2), 159–167.
https://doi.org/10.1207/s15327663jcp1401&2_18

RATIONALITY, PRAGMATICS, AND SOURCES

- Mills, C. M. (2013). Knowing When to Doubt: Developing a Critical Stance When Learning From Others. *Developmental Psychology, 49*(3).
<https://doi.org/10.1037/a0029500>
- Mills, C. M., & Keil, F. C. (2005). The Development of Cynicism. *Psychological Science, 16*(5), 385–390. <https://doi.org/10.1111/j.0956-7976.2005.01545.x>
- Mills, C. M., & Keil, F. C. (2008). Children’s developing notions of (im)partiality. *Cognition, 107*(2), 528–551. <https://doi.org/10.1016/j.cognition.2007.11.003>
- Misyak, J. B., & Chater, N. (2014). Virtual bargaining: a theory of social decision-making. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1655), 20130487–20130487.
<https://doi.org/10.1098/rstb.2013.0487>
- Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences, 18*(10), 512–519.
<https://doi.org/10.1016/j.tics.2014.05.010>
- Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes Factors for common designs. (Version R package version 0.9.12-2). Retrieved from <https://CRAN.R-project.org/package=BayesFactor>
- Morey, R. D., Rouder, J. N., & Jamil, T. (2015). Reference Manual: Package “BayesFactor.” Retrieved from <https://cran.r-project.org/web/packages/BayesFactor/BayesFactor.pdf>
- Morier, D. M., & Borgida, E. (1984). The Conjunction Fallacy: A Task Specific Phenomenon? *Personality and Social Psychology Bulletin, 10*(2), 243–252.
<https://doi.org/10.1177/0146167284102010>

RATIONALITY, PRAGMATICS, AND SOURCES

- Moro, R. (2009). On the nature of the conjunction fallacy. *Synthese*, 171(1), 1–24.
<https://doi.org/10.1007/s11229-008-9377-8>
- Moxey, L. M. (2006). Effects of what is expected on the focussing properties of quantifiers: A test of the presupposition-denial account. *Journal of Memory and Language*, 55(3), 422–439. <https://doi.org/10.1016/j.jml.2006.05.006>
- Moxey, L. M., & Sanford, A. J. (1986). QUANTIFIERS AND FOCUS. *Journal of Semantics*, 5(3), 189–206. <https://doi.org/10.1093/jos/5.3.189>
- Moxey, L. M., & Sanford, A. J. (2000). Communicating quantities: a review of psycholinguistic evidence of how expressions determine perspectives. *Applied Cognitive Psychology*, 14(3), 237–255.
[https://doi.org/10.1002/\(SICI\)1099-0720\(200005/06\)14:3<237::AID-ACP641>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1099-0720(200005/06)14:3<237::AID-ACP641>3.0.CO;2-R)
- Nan, X. (2007). The Relative Persuasive Effect of Gain- versus Loss-Framed Messages: Exploring the Moderating Role of the Desirability of End-States. *Journalism & Mass Communication Quarterly*, 84(3), 509–524.
<https://doi.org/10.1177/107769900708400307>
- Newell, B. R., Lagnado, D. A., & Shanks, D. R. (2015). *Straight Choices: The Psychology of Decision Making*. Hove, England: Psychology Press.
- Nieuwland, M. S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language*, 63(3), 324–346.
<https://doi.org/10.1016/j.jml.2010.06.005>
- Noveck, I. A., & Bonnefond, M. (2011). Squib: A deflationary account of invited inferences. *Belgian Journal of Linguistics*, 25, 195–208.
<https://doi.org/10.1075/bjl.25.09nov>

RATIONALITY, PRAGMATICS, AND SOURCES

- Noveck, I. A., & Reboul, A. (2008). Experimental Pragmatics: a Gricean turn in the study of language. *Trends in Cognitive Sciences*, *12*(11), 425–431.
<https://doi.org/10.1016/j.tics.2008.07.009>
- Noveck, I. A., & Sperber, D. (2004). *Experimental Pragmatics*. London, England: Palgrave Macmillan.
- Nyhan, B., & Reifler, J. (2010). When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior*, *32*(2), 303–330.
<https://doi.org/10.1007/s11109-010-9112-2>
- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608–631.
<https://doi.org/10.1037/0033-295X.101.4.608>
- Oaksford, M., & Chater, N. (2007). *Bayesian rationality: the probabilistic approach to human reasoning*. Oxford, England: Oxford University Press.
- Oaksford, M., & Chater, N. (2009a). Précis of Bayesian Rationality: The Probabilistic Approach to Human Reasoning. *Behavioral and Brain Sciences*, *32*(1), 69–84. <https://doi.org/10.1017/S0140525X09000284>
- Oaksford, M., & Chater, N. (2009b). The uncertain reasoner: Bayes, logic, and rationality. *Behavioral and Brain Sciences*, *32*(1), 105–120.
<https://doi.org/10.1017/S0140525X0900051X>
- Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(4), 883–899. <https://doi.org/10.1037/0278-7393.26.4.883>
- Oberauer, K., & Wilhelm, O. (2003). The meaning(s) of conditionals: Conditional probabilities, mental models, and personal utilities. *Journal of Experimental*

RATIONALITY, PRAGMATICS, AND SOURCES

Psychology: Learning, Memory, and Cognition, 29(4), 680–693.

<https://doi.org/10.1037/0278-7393.29.4.680>

O'Connor, D. B., Ferguson, E., & O'Connor, R. C. (2005). Intentions to use hormonal male contraception: The role of message framing, attitudes and stress appraisals. *British Journal of Psychology*, 96(3), 351–369.

<https://doi.org/10.1348/000712605X49114>

Ohm, E., & Thompson, V. A. (2004). Everyday reasoning with inducements and advice. *Thinking & Reasoning*, 10(3), 241–272.

<https://doi.org/10.1080/13546780342000061>

O'Keefe, D. J. (2011). Generalizing about the Persuasive Effects of Message Variations: The Case of Gain-Framed and Loss-Framed Appeals. In T. van Haaften, H. Jansen, J. De Jong, & W. Koetsenruijter (Eds.), *Bending Opinion: Essays on Persuasion in the Public Comain* (pp. 117–132). Leiden: Leiden University Press.

O'Keefe, D. J. (2012). The argumentative structure of some persuasive appeal variations. In F. H. van Eemeren & B. J. Garssen (Eds.), *Topical themes in argumentation: Twenty exploratory studies* (pp. 291–306). New York, NY: Springer.

O'Keefe, D. J. (2013). The relative persuasiveness of different forms of arguments-from-consequences: A review and integration. In E. L. Cohen (Ed.), *Communication Yearbook 36* (pp. 109–135). New York, NY: Routledge.

O'Keefe, D. J., & Jensen, J. D. (2006). The Advantages of Compliance or the Disadvantages of Noncompliance? A Meta-Analytic Review of the Relative Persuasive Effectiveness of Gain-Framed and Loss-Framed Messages.

Annals of the International Communication Association, 30(1), 1–43.

<https://doi.org/10.1080/23808985.2006.11679054>

O’Keefe, D. J., & Jensen, J. D. (2007). The Relative Persuasiveness of Gain-Framed Loss-Framed Messages for Encouraging Disease Prevention Behaviors: A Meta-Analytic Review. *Journal of Health Communication*, 12(7), 623–644.
<https://doi.org/10.1080/10810730701615198>

O’Keefe, D. J., & Jensen, J. D. (2009). The Relative Persuasiveness of Gain-Framed and Loss-Framed Messages for Encouraging Disease Detection Behaviors: A Meta-Analytic Review. *Journal of Communication*, 59(2), 296–316.
<https://doi.org/10.1111/j.1460-2466.2009.01417.x>

Olsson, E. J. (2005). *Against coherence: Truth, probability, and justification*. Oxford, England: Oxford University Press.

Olsson, E. J. (2011). A Simulation Approach to Veritistic Social Epistemology. *Episteme*, 8(2), 127–143. <https://doi.org/10.3366/epi.2011.0012>

Olsson, E. J., & Schubert, S. (2007). Reliability conducive measures of coherence. *Synthese*, 157(3), 297–308. <https://doi.org/10.1007/s11229-006-9056-6>

Olsson, E. J., & Vallinder, A. (2013). Norms of assertion and communication in social networks. *Synthese*, 190(13), 2557–2571.
<https://doi.org/10.1007/s11229-013-0313-1>

Orbell, S., & Hagger, M. (2006). Temporal framing and the decision to take part in type 2 diabetes screening: Effects of individual differences in consideration of future consequences on persuasion. *Health Psychology*, 25(4), 537–548.
<https://doi.org/10.1037/0278-6133.25.4.537>

RATIONALITY, PRAGMATICS, AND SOURCES

- Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review*, *62*(1), 42–55.
<https://doi.org/10.1037/h0048153>
- Over, D. E. (2016). The paradigm shift in the psychology of reasoning. In L. Macchi, M. Bagassi, & R. Viale (Eds.), *Cognitive Unconscious and Human Rationality* (pp. 79–98). Cambridge, MA: MIT Press.
- Over, D. E., Hadjichristidis, C., Evans, J. S. B. T., Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology*, *54*(1), 62–97. <https://doi.org/10.1016/j.cogpsych.2006.05.002>
- Paolacci, G., & Chandler, J. (2014). Inside the Turk Understanding Mechanical Turk as a Participant Pool. *Current Directions in Psychological Science*, *23*(3), 184–188. <https://doi.org/10.1177/0963721414531598>
- Parikh, P. (2001). *The Use of Language*. Palo Alto: CSLI Publications.
- Paterson, K. B., Sanford, A. J., Moxey, L. M., & Dawydiak, E. (1998). Quantifier Polarity and Referential Focus during Reading. *Journal of Memory and Language*, *39*(2), 290–306. <https://doi.org/10.1006/jmla.1998.2561>
- Perelman, C., & Olbrechts-Tyteca, L. (1969). *The New Rhetoric: A Treatise on Argumentation*. (J. Wilkinson & P. Weaver, Trans.). Notre Dame: University of Notre Dame Press.
- Perfors, A., Tenenbaum, J. B., & Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition*, *118*(3), 306–338.
<https://doi.org/10.1016/j.cognition.2010.11.001>
- Peterson, M. (2009). *An Introduction to Decision Theory*. Cambridge, England: Cambridge University Press.

RATIONALITY, PRAGMATICS, AND SOURCES

- Petty, R. E., & Briñol, P. (2008). Persuasion: From Single to Multiple to Metacognitive Processes. *Perspectives on Psychological Science*, 3(2), 137–147. <https://doi.org/10.1111/j.1745-6916.2008.00071.x>
- Petty, R. E., & Cacioppo, J. T. (1984). Source Factors and the Elaboration Likelihood Model of Persuasion. *Advances in Consumer Research*, 11, 668–672.
- Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 19, pp. 123–205). Cambridge, MA: Academic Press.
Retrieved from
<http://www.sciencedirect.com/science/article/pii/S0065260108602142>
- Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, 41(5), 847–855. <https://doi.org/10.1037/0022-3514.41.5.847>
- Pfeifer, N., & Kleiter, G. D. (2010). The conditional in mental probability logic. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thought*. (pp. 145–166). Oxford, England: Oxford University Press.
- Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, 72(3), 346–354. <https://doi.org/10.1037/h0023653>
- Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech. *Proceedings of the National Academy of Sciences*, 105(3), 833–838. <https://doi.org/10.1073/pnas.0707192105>

RATIONALITY, PRAGMATICS, AND SOURCES

- Politzer, G., & Noveck, I. A. (1991). Are conjunction rule violations the result of conversational rule violations? *Journal of Psycholinguistic Research*, *20*(2), 83–103. <https://doi.org/10.1007/BF01067877>
- Politzer, G., Over, D. E., & Baratgin, J. (2010). Betting on conditionals. *Thinking & Reasoning*, *16*(3), 172–197. <https://doi.org/10.1080/13546783.2010.504581>
- Popper, K., & Miller, D. (1983). A proof of the impossibility of inductive probability. *Nature*, *302*(5910), 687–688. <https://doi.org/10.1038/302687a0>
- Portner, Paul. (2009). *Modality*. Oxford, England: Oxford University Press.
- Poulin-Dubois, D., Brooker, I., & Polonia, A. (2011). Infants prefer to imitate a reliable person. *Infant Behavior and Development*, *34*(2), 303–309. <https://doi.org/10.1016/j.infbeh.2011.01.006>
- Poulin-Dubois, D., & Chow, V. (2009). The effect of a looker's past reliability on infants' reasoning about beliefs. *Developmental Psychology*, *45*(6), 1576–1582. <https://doi.org/10.1037/a0016715>
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to Deal with “The Language-as-Fixed-Effect Fallacy”: Common Misconceptions and Alternative Solutions. *Journal of Memory and Language*, *41*(3), 416–426. <https://doi.org/10.1006/jmla.1999.2650>
- Read, D. (2007). Experienced utility: Utility theory from Jeremy Bentham to Daniel Kahneman. *Thinking & Reasoning*, *13*(1), 45–61. <https://doi.org/10.1080/13546780600872627>

RATIONALITY, PRAGMATICS, AND SOURCES

- Recanati, F. (2002). Unarticulated Constituents. *Linguistics and Philosophy*, 25(3), 299–345. <https://doi.org/10.1023/A:1015267930510>
- Recanati, F. (2004). *Literal Meaning*. Cambridge, England: Cambridge University Press. Retrieved from https://books.google.co.uk/books/about/Literal_Meaning.html?id=ItPyx3frII
- Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology*, 96(3), 538–558. <https://doi.org/10.1037/a0014038>
- Rips, L. J. (1994). *The Psychology of Proof: Deductive Reasoning in Human Thinking*. Cambridge, MA: MIT Press.
- Roberts, C. (1996). Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. *OSU Working Papers in Linguistics*, 49, 91–136.
- Robinson, E. J., Champion, H., & Mitchell, P. (1999). Children's ability to infer utterance veracity from speaker informedness. *Developmental Psychology*, 35(2), 535–546. <https://doi.org/10.1037/0012-1649.35.2.535>
- Rosenkrantz, R. . (1992). The justification of induction. *Philosophy of Science*, 15, 527–539. <https://doi.org/doi:10.1086/289693>
- Rothman, A. J., & Salovey, P. (1997). Shaping perceptions to motivate healthy behavior: The role of message framing. *Psychological Bulletin*, 121(1), 3–19. <https://doi.org/10.1037/0033-2909.121.1.3>
- Rothman, A. J., Salovey, P., Antone, C., Keough, K., & Martin, C. D. (1993). The Influence of Message Framing on Intentions to Perform Health Behaviors.

RATIONALITY, PRAGMATICS, AND SOURCES

Journal of Experimental Social Psychology, 29(5), 408–433.

<https://doi.org/10.1006/jesp.1993.1019>

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: links between preschoolers' theory of mind and semantic development. *Child Development*, 72(4), 1054–1070.

Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, and knowledge. In P. C. Wason & P. N. Johnson-Laird (Eds.), *Thinking: Readings in Cognitive Science* (pp. 421–435). Cambridge: Cambridge University Press.

Schellens, P. J., & De Jong, M. (2004). Argumentation Schemes in Persuasive Brochures. *Argumentation*, 18(3), 295–323.

<https://doi.org/10.1023/B:ARGU.0000046707.68172.35>

Schiffer, S. R. (1972). *Meaning*. Oxford, England: Oxford University Press.

Schneider, T. R., Salovey, P., Apanovitch, A. M., Pizarro, J., McCarthy, D., Zullo, J., & Rothman, A. J. (2001). The effects of message framing and ethnic targeting on mammography use among low-income women. *Health Psychology*, 20(4), 256–266. <https://doi.org/10.1037/0278-6133.20.4.256>

Schreier, M., Groeben, N., & Christmann, U. (1995). “That’s not fair!” argumentational integrity as an ethics of argumentative communication. *Argumentation*, 9(2), 267–289. <https://doi.org/10.1007/BF00721962>

Schroyens, W. (2010). A meta-analytic review of thinking about what is true, possible, and irrelevant in reasoning from or reasoning about conditional propositions. *European Journal of Cognitive Psychology*, 22(6), 897–921. <https://doi.org/10.1080/09541440902928915>

RATIONALITY, PRAGMATICS, AND SOURCES

- Schubert, S. (2012). Coherence reasoning and reliability: a defense of the Shogenji measure. *Synthese*, 187(2), 305–319. <https://doi.org/10.1007/s11229-010-9856-6>
- Schum, D. A. (1981). Sorting out the effects of witness sensitivity and response-criterion placement upon the inferential value of testimonial evidence. *Organizational Behavior and Human Performance*, 27(2). [https://doi.org/10.1016/0030-5073\(81\)90045-3](https://doi.org/10.1016/0030-5073(81)90045-3)
- Schwarz, N. (1996). *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Schwarz, N., Knauper, B., Hippler, H.-J., Noelle-Neumann, E., & Clark, L. (1991). Rating Scales: Numeric Values May Change the Meaning of Scale Labels. *Public Opinion Quarterly*, 55(4), 570. <https://doi.org/10.1086/269282>
- Schwarz, N., Sanna, L. J., Skurnik, I., & Yoon, C. (2007). Metacognitive Experiences and the Intricacies of Setting People Straight: Implications for Debiasing and Public Information Campaigns. *Advances in Experimental Social Psychology*, 39, 127–161. [https://doi.org/10.1016/S0065-2601\(06\)39003-X](https://doi.org/10.1016/S0065-2601(06)39003-X)
- Schwarz, N., Strack, F., Hilton, D., & Naderer, G. (1991). Base Rates, Representativeness, and the Logic of Conversation: The Contextual Relevance of “Irrelevant” Information. *Social Cognition*, 9(1), 67–84. <https://doi.org/10.1521/soco.1991.9.1.67>
- Searle, J. S. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge, England: Cambridge University Press.

RATIONALITY, PRAGMATICS, AND SOURCES

- Shafto, P., Eaves, B., Navarro, D. J., & Perfors, A. (2012). Epistemic trust: modeling children's reasoning about others' knowledge and intent. *Developmental Science*, *15*(3), 436–447. <https://doi.org/10.1111/j.1467-7687.2012.01135.x>
- Sher, S., & McKenzie, C. R. M. (2006). Information leakage from logically equivalent frames. *Cognition*, *101*(3), 467–494. <https://doi.org/10.1016/j.cognition.2005.11.001>
- Sher, S., & McKenzie, C. R. M. (2014). Options as information: Rational reversals of evaluation and preference. *Journal of Experimental Psychology: General*, *143*(3), 1127–1143. <https://doi.org/10.1037/a0035128>
- Sides, A., Osherson, D., Bonini, N., & Viale, R. (2002). On the reality of the conjunction fallacy. *Memory & Cognition*, *30*(2), 191–198. <https://doi.org/10.3758/BF03195280>
- Simons, M. (2007). Observations on embedding verbs, evidentiality, and presupposition. *Lingua*, *117*(6), 1034–1056. <https://doi.org/10.1016/j.lingua.2006.05.006>
- Singmann, H., Klauer, K. C., & Beller, S. (2016). Probabilistic conditional reasoning: Disentangling form and content with the dual-source model. *Cognitive Psychology*, *88*, 61–87. <https://doi.org/10.1016/j.cogpsych.2016.06.005>
- Skovgaard-Olsen, N. (2016). Motivating the Relevance Approach to Conditionals. *Mind & Language*, *31*(5), 555–579. <https://doi.org/10.1111/mila.12120>
- Skovgaard-Olsen, N., Kellen, D., Krahl, H., & Klauer, K. C. (2017). *Relevance differently affects the truth, acceptability, and probability evaluations of “and”, “but”, “therefore”, and “if then.”* Manuscript submitted for publication.

RATIONALITY, PRAGMATICS, AND SOURCES

Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2016a). Relevance and Reason Relations. *Cognitive Science*, 1202–1215.

<https://doi.org/10.1111/cogs.12462>

Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2016b). The relevance effect and conditionals. *Cognition*, 150, 26–36.

<https://doi.org/10.1016/j.cognition.2015.12.017>

Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How Warnings about False Claims Become Recommendations. *Journal of Consumer Research*, 31(4), 713–724. <https://doi.org/10.1086/426605>

Slater, M. D., & Rouner, D. (1996). How Message Evaluation and Source Attributes May Influence Credibility Assessment and Belief Change. *Journalism & Mass Communication Quarterly*, 73(4), 974–991.

<https://doi.org/10.1177/107769909607300415>

Smith, V. L., & Ellsworth, P. C. (1987). The social psychology of eyewitness accuracy: Misleading questions and communicator expertise. *Journal of Applied Psychology*, 72(2), 294–300. <https://doi.org/10.1037/0021-9010.72.2.294>

Snizek, J. A. (1989). An examination of group process in judgmental forecasting. *International Journal of Forecasting*, 5(2), 171–178.

[https://doi.org/10.1016/0169-2070\(89\)90084-8](https://doi.org/10.1016/0169-2070(89)90084-8)

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, 57(1), 31–95.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010). Epistemic Vigilance. *Mind & Language*, 25(4), 359–393.

<https://doi.org/10.1111/j.1468-0017.2010.01394.x>

RATIONALITY, PRAGMATICS, AND SOURCES

- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition* (Second Edition). Oxford, England: Blackwell.
- Sperber, D., & Wilson, D. (2002). Pragmatics, Modularity and Mind-reading. *Mind & Language*, 17(1–2), 3–23. <https://doi.org/10.1111/1468-0017.00186>
- Stalnaker, R. (1975). Indicative conditionals. *Philosophia*, 5(3), 269–286. <https://doi.org/10.1007/BF02379021>
- Stanovich, K. ., & West, R. F. (1999). Discrepancies Between Normative and Descriptive Models of Decision Making and the Understanding/Acceptance Principle. *Cognitive Psychology*, 38(3), 349–385. <https://doi.org/10.1006/cogp.1998.0700>
- Stanovich, K. E. (2005). *The Robot's Rebellion: Finding Meaning in the Age of Darwin*. Chicago, IL: University of Chicago Press.
- Stanovich, K. E. (2012). On the distinction between rationality and intelligence: implications for understanding individual differences in reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning* (pp. 343–365). Oxford, England: Oxford University Press.
- Stanovich, K. E. (2016). The Comprehensive Assessment of Rational Thinking. *Educational Psychologist*, 51(1), 23–34. <https://doi.org/10.1080/00461520.2015.1125787>
- 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.
- Stanovich, K. E., West, R. F., & Toplak, M. E. (2016). *The Rationality Quotient: Toward a Test of Rational Thinking*. Cambridge, MA: MIT Press.

RATIONALITY, PRAGMATICS, AND SOURCES

- Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, 48(6), 1467–1478. <https://doi.org/10.1037/0022-3514.48.6.1467>
- Stasser, G., & Titus, W. (2003). Hidden Profiles: A Brief History. *Psychological Inquiry*, 14(3–4), 304–313. <https://doi.org/10.1080/1047840X.2003.9682897>
- Stein, E. (1996). *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*. Clarendon Press.
- Stenning, K., & Lambalgen, M. van. (2012). *Human Reasoning and Cognitive Science*. Cambridge, MA: MIT Press.
- Stevenson, R. J., & Over, D. E. (2001). Reasoning from uncertain premises: Effects of expertise and conversational context. *Thinking & Reasoning*, 7(4), 367–390. <https://doi.org/10.1080/13546780143000080>
- Stewart, A. J., Haigh, M., & Ferguson, H. J. (2013). Sensitivity to speaker control in the online comprehension of conditional tips and promises: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(4), 1022–1036. <https://doi.org/10.1037/a0031513>
- Strack, F., Schwarz, N., & Wänke, M. (1991). Semantic and Pragmatic Aspects of Context Effects in Social and Psychological Research. *Social Cognition*, 9(1), 111–125. <https://doi.org/10.1521/soco.1991.9.1.111>
- Strawson, P. F. (1964). Intention and Convention in Speech Acts. *The Philosophical Review*, 73(4), 439–460. <https://doi.org/10.2307/2183301>
- Sullivan, M. (2016). Tightening the Screws on Anonymous Sources. Retrieved June 15, 2017, from <https://publiceditor.blogs.nytimes.com/2016/03/15/new-york-times-anonymous-sources-policy-public-editor/>

RATIONALITY, PRAGMATICS, AND SOURCES

- Surian, L., Baron-Cohen, S., & Van der Lely, H. (1996). Are Children with Autism Deaf to Gricean Maxims? *Cognitive Neuropsychiatry*, *1*(1), 55–72.
<https://doi.org/10.1080/135468096396703>
- Surowiecki, J. (2005). *The Wisdom of Crowds*. New York, NY: Knopf Doubleday Publishing Group.
- Swann, W. B., Giuliano, T., & Wegner, D. M. (1982). Where leading questions can lead: The power of conjecture in social interaction. *Journal of Personality and Social Psychology*, *42*(6), 1025–1035. <https://doi.org/10.1037/0022-3514.42.6.1025>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics*. London, England: Pearson.
- Taber, C. S., & Lodge, M. (2006). Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science*, *50*(3), 755–769.
<https://doi.org/10.1111/j.1540-5907.2006.00214.x>
- Tentori, K., & Crupi, V. (2012). On the conjunction fallacy and the meaning of and, yet again: A reply to. *Cognition*, *122*(2), 123–134.
<https://doi.org/10.1016/j.cognition.2011.09.002>
- Tentori, K., Crupi, V., & Russo, S. (2013). On the determinants of the conjunction fallacy: Probability versus inductive confirmation. *Journal of Experimental Psychology: General*, *142*(1), 235–255. <https://doi.org/10.1037/a0028770>
- Tetlock, P. E., Mellers, B. A., Rohrbaugh, N., & Chen, E. (2014). Forecasting Tournaments: Tools for Increasing Transparency and Improving the Quality of Debate. *Current Directions in Psychological Science*, *23*(4), 290–295.
<https://doi.org/10.1177/0963721414534257>

RATIONALITY, PRAGMATICS, AND SOURCES

- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. (New International Edition). London, England: Penguin.
- Thompson, V. A. (2010). Towards a metacognitive dual process theory of conditional reasoning. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: probability and logic in human thinking* (pp. 335–354). Oxford, England: Oxford University Press.
- Tonhauser, J. (2012). Diagnosing (not-)at-issue content. In *In Proceedings of Semantics of Under-represented Languages of the Americas (SULA)* (Vol. 6, pp. 239–254). UMass, Amherst: GSLA.
- Toulmin, S. (1958). *The Uses of Argument*. Cambridge, England: Cambridge University Press.
- Toulmin, S., Rieke, R., & Janik, A. (1979). *An introduction to reasoning*. New York NY: Macmillan.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science (New York, N.Y.)*, *211*(4481), 453–458.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*(4), 293–315. <https://doi.org/10.1037/0033-295X.90.4.293>
- Tweney, R. D., Doherty, M. E., & Kleiter, G. D. (2010). The pseudodiagnosticity trap: Should participants consider alternative hypotheses? *Thinking & Reasoning*, *16*(4), 332–345. <https://doi.org/10.1080/13546783.2010.525860>
- Uskul, A. K., Sherman, D. K., & Fitzgibbon, J. (2009). The cultural congruency effect: Culture, regulatory focus, and the effectiveness of gain- vs. loss-

- framed health messages. *Journal of Experimental Social Psychology*, 45(3), 535–541. <https://doi.org/10.1016/j.jesp.2008.12.005>
- van Assema, P., Martens, M., Ruiter, R. A., & Brug, J. (2001). Framing of nutrition education messages in persuading consumers of the advantages of a healthy diet. *Journal of Human Nutrition and Dietetics: The Official Journal of the British Dietetic Association*, 14(6), 435–442.
- Van Der Auwera, J. (1996). Modality: The Three-layered Scalar Square. *Journal of Semantics*, 13(3), 181–195. <https://doi.org/10.1093/jos/13.3.181>
- Van der Henst, J.-B., & Sperber, D. (2004). Testing the cognitive and communicative principles of relevance. In I. A. Noveck & D. Sperber (Eds.), *Experimental Pragmatics* (pp. 229–279). London, England: Palgrave.
- Verstraete, J.-C. (2005). Scalar quantity implicatures and the interpretation of modality. *Journal of Pragmatics*, 37(9), 1401–1418. <https://doi.org/10.1016/j.pragma.2005.02.003>
- Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and Opportunities in Nonverbal and Verbal Lie Detection. *Psychological Science in the Public Interest*, 11(3), 89–121. <https://doi.org/10.1177/1529100610390861>
- Wallin, A., Paradis, C., & Katsikopoulos, K. V. (2016). Evaluative polarity words in risky choice framing. *Journal of Pragmatics*, 106, 20–38. <https://doi.org/10.1016/j.pragma.2016.09.005>
- Walton, D. N. (2003). *Relevance in Argumentation*. Routledge.
- Walton, D. N. (2008). *Informal Logic: A Pragmatic Approach* (2nd Edition). Cambridge, England: Cambridge University Press.
- Walton, D. N., & Macagno, F. (2016). Profiles of Dialogue for Relevance. *Informal Logic*, 36(4), 523–562. <https://doi.org/10.22329/il.v36i4.4586>

RATIONALITY, PRAGMATICS, AND SOURCES

- Walton, D. N., Reed, C., & Macagno, F. (2008). *Argumentation Schemes*. Cambridge, England: Cambridge University Press.
- Wason, P. C. (1965). The contexts of plausible denial. *Journal of Verbal Learning & Verbal Behavior*, 4(1), 7–11. [https://doi.org/10.1016/S0022-5371\(65\)80060-3](https://doi.org/10.1016/S0022-5371(65)80060-3)
- Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.) (Vol. 1). London, England: Penguin.
- Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, 67(6), 1049–1062. <https://doi.org/10.1037/0022-3514.67.6.1049>
- Wegener, D. T., Petty, R. E., & Klein, D. J. (1994). Effects of mood on high elaboration attitude change: The mediating role of likelihood judgments. *European Journal of Social Psychology*, 24(1), 25–43. <https://doi.org/10.1002/ejsp.2420240103>
- Wilcox, R. R. (2016). *Understanding and Applying Basic Statistical Methods Using R*. Hoboken, N.J.: John Wiley & Sons.
- Wilson, D., & Carston, . (2007). A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.), *Advances in Pragmatics* (pp. 230–260). London, England: Palgrave.
- Wolford, G., Taylor, H. A., & Beck, J. R. (1990). The conjunction fallacy? *Memory & Cognition*, 18(1), 47–53. <https://doi.org/10.3758/BF03202645>
- Wright, E. F., & Wells, G. L. (1988). Is the Attitude-Attribution Paradigm Suitable for Investigating the Dispositional Bias? *Personality and Social Psychology Bulletin*, 14(1), 183–190. <https://doi.org/10.1177/0146167288141018>
- Zenker, F. (2012). *Bayesian Argumentation: The practical side of probability*. New York, NY: Springer Science & Business Media.

