University College London

Doctoral Thesis

Constructing the world: Active causal learning in cognition

Author:
Neil Robert Bramley

Supervisors:
Prof. David A. Lagnado
Prof. Peter Dayan

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

Department of Experimental Psychology

February 2017

Declaration of Authorship

I, Neil Robert Bramley, declare that this thesis titled, "Constructing the world: Active causal learning in cognition" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at University College London.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Date: _____

UNIVERSITY COLLEGE LONDON

Abstract

Brain Sciences
Department of Experimental Psychology

Doctor of Philosophy

Constructing the world: Active causal learning in cognition

by Neil Robert Bramley

Humans are adept at constructing causal models of the world that can support prediction, explanation, simulation-based reasoning, planning and control. In this thesis I explore how people learn about the causal world interacting with it, and how they represent and modify their causal knowledge as they gather evidence. Over 10 experiments and modelling, I show that interventional and temporal cues, along with top-down hierarchical constraints, inform the gradual evolution and adaptation of increasingly rich causal representations.

Chapters 1 and 2 develop a rational analysis of the problems of learning and representing causal structure, and choosing *interventions*, that perturb the world in ways that reveal its structure. Chapters 3–5 focus on structure learning over sequences of discrete trials, in which learners can intervene by setting variables within a causal system and observe the consequences. The second half of the thesis generalises beyond the discrete trial learning case, exploring interventional causal learning in situations where events occur in continuous time (Chapters 6 and 7); and in spatiotemporally rich physical "microworlds" (Chapter 8). Throughout the experiments, I find that both children and adults are robust active causal learners, able to deal with noise and complexity even as normative judgment and intervention selection become radically intractable. To explain their success, I develop scalable process level accounts of both causal structure learning and intervention selection inspired by approximation algorithms in machine learning. I show that my models can better explain patterns of behaviour than a range of alternatives as well as shedding light on the source of common biases including confirmatory testing, anchoring effects and probability matching. Finally, I propose a close relationship between active learning and active aspects of cognition including thinking, decision making and executive control.

Acknowledgements

This thesis would not exist without the fantastic supervision of David Lagnado. Dave's supervisory credentials are evident in his growing collection of "World's best supervisor" mugs and I am already planning the latest addition. Throughout MSc, MRes and PhD, Dave's generous, intelligent and above-all supportive supervision kept me on track and gave me confidence and motivation to develop my research. He was there for every stage of my development giving me the nudge, or the back-up I needed to run my first experiment, program my first model, talk at my first conference, give my first lecture, publish my first paper, write my first review and my first grant application, and host my first symposium. Along the way, Dave's sharp eye for the important questions and refreshing irreverence for the nastier parts of academia have made working with him a privilege and a pleasure.

My girlfriend Paula Parpart deserves all the thanks in the world, for her love, and for putting up with me while she balances lecturing, business, music and her own PhD. Arriving at UCL in 2010 with an arts degree, I would not have made it far without Maarten Speekenbrink who taught me statistics, helped me develop my first models and has been a friend and collaborator ever since. I also owe a debt to my second supervisor Peter Dayan, whose formidable intellect, expertise and attention to detail, combined with amazing patience helped me take my work to a much higher level. Tobias Gerstenberg has also been a great friend and an unofficial third supervisor, teaching me the tricks of the trade as we collaborated on various projects (three featuring in this thesis). Eric Schulz, a close friend who started at UCL the same day as me, has led the way ever since and has earned his own sentence whether he likes it or not.

Thanks to Philip Treleaven, Yonita Carter and the Financial Computing DTC (and its parent organisation the EPSRC) for funding me. Thanks to Tom Griffiths (Berkeley) and Todd Gureckis (NYU), who hosted my research visit to the USA and whose work continues to inspire me. Thanks to Josh Abbott, Jessica Hamrick, Falk Lieder and Mike Pacer for making me feel welcome when I visited Berkeley, and Anna Coenen, Alex Rich and Anselm Rothe for doing the same at NYU (where I am headed next). Thanks to the many other academics I have been lucky enough to collaborate with while at UCL: Vincenzo Crupi (Turin), Caren Frosch (Leicester), Ralf Mayrhofer (Göttingen), Teresa McCormack (Belfast), Björn Meder (Max Planck, Berlin), Johnathan Nelson (Max Planck, Berlin), Hansjörg Neth (Konstanz), Azzura Ruggeri (Max Planck, Berlin) and Joshua Tenenbaum (MIT). Going further back, thanks to David Bain, Chris Lindsay and Susan Stuart (Glasgow) for sparking my interest in cognition.

Not least, I want to thank the academics, lab mates, friends and peers (past and present) who have made UCL a fantastic place to work: Christos Bechlivanidis, Chloe Bukata, Leo Cohen, Will Crichton, George Deane, Saoirse Connor Desai, John Draper, Neele Engelmann, Moritz Haller, Tom Hardwicke, Adam Harris, Nigel Harvey, Pia Horbacki, Anne Hsu, Sarah Jenkins, Lara Kirfel, Manos Konstantinidis, Brad Love, Clare Machan, Jens Koed Madsen, Simone Malejka, Laura de Molière, Phil Newell (who organises a great workshop), Milena Nikolic, Keith O'Brien, Magda Osman, Fiona Patrick, Philip Parnämets, Toby Pilditch, Costi Rezlescu, David Shanks, Sarah Smith, Paul Troop, Pablo Léon Víllagra, Miguel Vadillo, Marion Vorms, Laura Weis, Jacob de Zoete, and anyone else I have forgotten.

Finally, of course none of this would have been possible without the unerring love and support of my parents Alison and Glen.

Contents

Declaration of Authorship **2**

Abstract **3**

Acknowledgements **4**

Contents **5**

List of Figures **11**

List of Tables **15**

Abbreviations **17**

Symbols **19**

1 Introduction **27**

2 Causal cognition **33**
2.1 Computationalism 33
2.2 Rational analysis 35
2.3 Bayesian inference 35
2.4 Probabilistic graphical models 38
2.5 Causal Bayesian networks 43
2.5.1 Intervention 46
2.6 Active learning 47
2.6.1 Optimal intervening 50
2.7 Intractability 52
2.7.1 Heuristics 52
2.7.2 Rational process models 54
2.7.3 Bounded intervention selection 60
2.8 Beyond Bayesian networks 62
2.8.1 Richer functional forms 63
2.8.2 Intuitive theories, and hierarchical representations 63
2.8.3 The role of time 66

2.8.4 The role of space . . . 67
2.8.5 Even more approximation . . . 69
2.9 Summary . . . 69

3 Learning causal structure through intervention 71
3.1 The Causal Bayesian network framework . . . 73
3.1.1 Representation . . . 73
3.1.2 Inference . . . 74
3.1.3 Choosing interventions . . . 76
3.2 Comparing intervention objectives . . . 79
3.3 Experiment 1: Learning structure by intervention . . . 81
3.4 Experiment 2: Information button and summary . . . 86
3.4.1 Intervention choices . . . 88
3.5 Modelling intervention selection and causal judgments . . . 94
3.5.1 Expectancy-based models . . . 97
3.5.2 Heuristic models . . . 103
3.5.3 Model estimation and comparison . . . 105
3.5.4 Model fit results . . . 105
3.6 Discussion . . . 110
3.7 Conclusions . . . 113

4 Children's active causal learning 115
4.1 Black box paradigm . . . 119
4.2 Experiment 3: Children's interventional learning . . . 122
4.2.1 Method . . . 122
4.2.2 Results . . . 125
4.2.3 Modelling Interventions . . . 127
4.2.4 Discussion . . . 131
4.3 General Discussion . . . 135
4.4 Conclusions . . . 136

5 Scaling up 137
5.1 Behavioural patterns and existing explanations . . . 140
5.2 Algorithms for causal learning with limited resources . . . 142
5.2.1 Approximating with a few hypotheses . . . 143
5.2.2 Sequential local search . . . 143
5.3 Neurath's ship: An algorithmic-level model of sequential belief change . . . 145
5.3.1 Resampling, hill climbing or random change . . . 146
5.3.2 Search length . . . 146
5.3.3 Putting these together . . . 147
5.4 Selecting interventions on Neurath's ship: A local uncertainty schema . . . 147
5.4.1 The two stages of the schema . . . 149
5.4.2 Three varieties of local focus . . . 151
5.4.3 Implications of the schema . . . 153
5.5 Comparing model predictions to experiments . . . 153
5.5.1 Bramley, Lagnado & Speekenbrink (2015) . . . 155
5.5.2 Motivating the new experiments . . . 159

5.6 Experiment 4: Learning larger causal models . . . 160
5.6.1 Methods . . . 160
5.6.2 Results and discussion . . . 163
5.7 Experiment 5: Unknown strengths . . . 166
5.7.1 Methods . . . 168
5.7.2 Results and discussion . . . 170
5.7.3 Summary of Experiments . . . 176
5.8 Modelling individual behaviour . . . 177
5.8.1 Judgments . . . 178
5.8.2 Interventions . . . 183
5.9 General Discussion . . . 186
5.9.1 Limitations of Neurath's ship . . . 188
5.9.2 Alternative approximations and representations . . . 192
5.9.3 Choosing interventions aboard Neurath's ship . . . 193
5.9.4 The navy of one . . . 196
5.9.5 Scope of the theory . . . 198
5.10 Conclusions . . . 199

6 The role of time in causal learning 201
6.1 Existing research . . . 203
6.2 Modelling causal induction from temporal information . . . 207
6.2.1 The learning problem . . . 208
6.2.2 Bayesian models of learning . . . 208
6.2.3 Only order matters . . . 210
6.2.4 Timing matters . . . 212
6.2.5 Summary . . . 218
6.3 Overview of Experiments . . . 219
6.3.1 The task . . . 219
6.4 Experiment 6: One-shot inferences . . . 221
6.4.1 Methods . . . 221
6.4.2 Results . . . 223
6.4.3 Discussion . . . 225
6.5 Experiment 7: Integrating evidence . . . 226
6.5.1 Methods . . . 227
6.5.2 Results . . . 230
6.5.3 Discussion . . . 235
6.6 Experiment 8: Learning from timing variability alone . . . 236
6.6.1 Methods . . . 237
6.6.2 Results . . . 241
6.6.3 Discussion: Sensitivity to timing . . . 244
6.7 General Discussion . . . 248
6.7.1 Nonsimultaneity and simultaneity . . . 249
6.7.2 Causal time perception . . . 249
6.7.3 Modality . . . 250
6.7.4 Conjunctive influence . . . 250
6.7.5 The blessing of variability . . . 251
6.7.6 Toward a process model . . . 252

6.7.7 Building richer causal representations . . . 252
6.8 Conclusions . . . 253

7 Intervening in time **255**
7.1 The learning problem . . . 256
7.1.1 Interventions as structuring learning in time . . . 258
7.1.2 Variability and positive testing . . . 258
7.1.3 Causal cycles . . . 259
7.2 Modelling continuous time causal learning . . . 260
7.2.1 Normative inference . . . 260
7.2.2 Incremental construction heuristics . . . 262
7.3 Experiment 9: Intervening in time . . . 265
7.3.1 Methods . . . 265
7.3.2 Results . . . 269
7.4 Modelling the judgments . . . 273
7.5 Discussion . . . 275
7.5.1 Looking backwards or forwards? . . . 276

8 Intervening in space **279**
8.1 Active learning in richer domains . . . 280
8.1.1 Intuitive physics . . . 281
8.2 Experiment 10: Intervening in space and time . . . 284
8.2.1 Methods . . . 285
8.2.2 Results . . . 288
8.2.3 Interim discussion . . . 290
8.2.4 Measuring information . . . 290
8.2.5 Natural experiments . . . 294
8.3 Discussion . . . 296
8.4 Conclusions . . . 299

9 General discussion **301**
9.1 Causal theory change reconsidered . . . 301
9.1.1 Getting started . . . 302
9.1.2 Navigating infinite hypothesis spaces . . . 303
9.1.3 Vertical structure . . . 304
9.2 Intervention reconsidered . . . 305
9.2.1 Optimal self teaching . . . 305
9.2.2 Interventions that explore versus interventions that test . . . 306
9.2.3 Intuitive theories of intervention — learning to actively learn . . . 308
9.2.4 From intervention to adaptive control . . . 308
9.3 Turning control inward . . . 309
9.4 Conclusions . . . 311

A Further formalizing Neurath's ship **313**
A.1 Representation and inference . . . 313
A.2 Intervention choice . . . 314

A.3 An algorithmic-level model of sequential belief change 315
A.4 A local uncertainty schema 317
A.5 Additional modelling details 318
A.6 Model recovery 320

B Representing time **321**
B.1 Collider likelihood 321
B.2 Simple Monte Carlo estimation: Experiment 6 and 7 323
B.3 Markov Chain Monte Carlo estimation: Experiment 8 324
B.4 Checking sensitivity to priors 324
B.5 Correlations Among Heuristic Predictors in Experiment 8 327
B.6 Estimating posteriors in Experiment 9 327

C Physics worlds **331**
C.1 Box2D settings 331
C.2 Online information measures 332

References **335**

List of Figures

2.1 Schematic of Active Causal Learning . . . 38
2.2 Graphical Model Examples . . . 40
2.3 An Example of Markov Equivalence . . . 43
2.4 All 25 Directed Acyclic Graphs on Three Variables . . . 44
2.5 Interventional Learning from Covariational Evidence . . . 51
2.6 Using Simple Monte Carlo Sampling to Estimate the Probability of a Connection Between X_1 and X_2 . . . 56
2.7 Two Examples of Sample Based Approximate Causal Inference . . . 58
2.8 What Can You Learn from Interventions in Deterministic Three Variable Problems? . . . 62
2.9 Causal Theories and Recursive Networks . . . 65

3.1 Causal Model Representation . . . 74
3.2 Example Value Differences . . . 80
3.3 Procedure for Experiments 1 and 2 . . . 84
3.4 Experiments 1 and 2: Histograms of Scores . . . 85
3.5 Experiment 2: Interface . . . 87
3.6 Experiments 1 and 2: Interventional Efficiency . . . 91
3.7 Experiments 1 and 2: Simple and Controlled Tests . . . 92
3.8 Experiments 1 and 2: Test structures and averaged judgments . . . 93
3.9 A Controlled Intervention . . . 94
3.10 Schematic of Relationships Between Models . . . 96
3.11 Flowchart of Expectancy Based Models . . . 98
3.12 Process Trees for the Simple Endorser . . . 104
3.13 Process Tree for the Disambiguator . . . 105
3.14 Visualisation of Models . . . 106
3.15 Mean Correlation between Intervention Values . . . 107

4.1 Black Box Paradigm for Developmental Experiments . . . 120
4.2 Experiment 3: Apparatus . . . 123
4.3 Experiment 3: Responses by Structure, Age and Trial Type . . . 126
4.4 Interventions Selected by a 6- to 7-year-old in the Fork Trial . . . 130

5.1 An Illustration of NS Model of Causal Belief Updating . . . 148
5.2 An Illustrative Example of Local Focused Uncertainty Minimisation . . . 154
5.3 Experiments 1 and 3 revisited: Performance and Interventions . . . 158
5.4 Experimental Procedure . . . 161
5.5 Experiment 4: True Models, Average Judgments and Posteriors . . . 164

5.6 Experiment 4: Performance and Interventions . . . 167
5.7 Experiment 5: Additional Measures . . . 170
5.8 Experiment 5: True Models, Average Performance and Posteriors . . . 170
5.9 Experiment 5: Priors and Strength Estimates . . . 171
5.10 Experiment 5: Performance and Interventions . . . 172
5.11 Experiment 5: Free Explanations and Agreed Codes . . . 176
5.12 Modelling Parameters . . . 186
5.13 Model–Free-Response Correspondence . . . 187

6.1 Experiments 6–8: Possible Causal Structures . . . 209
6.2 Order Based Likelihoods . . . 211
6.3 Order Based Model Predictions . . . 212
6.4 Four Example Gamma Distributions . . . 215
6.5 Delay Sensitive Models And Predictions . . . 217
6.6 Experiments 6–7: Inferface . . . 220
6.7 Experiment 6: Results . . . 222
6.8 Comparison of Fork, Chain and Collider for Clips 1–7 . . . 224
6.9 Experiment 7 Interface . . . 229
6.10 Experiment 7: Elicited Priors . . . 230
6.11 Experiment 7: Elicited Likelihoods . . . 231
6.12 Experiment 7: Posterior Judgments . . . 233
6.13 Experiment 8: Stimuli and Model Predictions . . . 238
6.14 Experiment 8 Interface . . . 240
6.15 Experiment 8: Judgments by Device . . . 242
6.16 Experiment 8: Judgment Models . . . 247

7.1 Examples of "Free Operant Intervention" . . . 257
7.2 Continuous Time Causal Learning Heuristics . . . 264
7.3 Generative models tested in Experiment 9 . . . 266
7.4 Delay distributions by condition. . . . 267
7.5 Experiment 9: Procedure . . . 269
7.6 Experiment 9: Accuracy by Condition . . . 270
7.7 Experiment 9: Intervention Choice Prevalence By Component And Accuracy By Edge . . . 271

8.1 Experiment 10: Schematic Display of Microworlds . . . 283
8.2 Eliciting judgments in Experiment 10. Participants answered two forced choice questions and set confidence sliders. . . . 287
8.3 Experiment 10: Performance by Condition . . . 289
8.4 Experiment 10: Confusion Matrices . . . 290
8.5 Calculating Model Likelihoods from Simulated Trajectories . . . 292
8.6 Experiment 10: Remaining Uncertainty for Active and Passive Participants 293
8.7 Experiment 10: Predicted Divergence Comparisons . . . 294
8.8 Experiment 10: Examples of Natural Experiments . . . 296
8.9 Timeline for an Active Trial Exemplifying *Shaking* . . . 297
8.10 Timeline for an Active Trial Exemplifying *Deconfounding* . . . 298

9.1 Three stages of intervention . . . 307

B.1 Experiment 8: Sensitivity of Model Predictions on α Prior 325
B.2 Experiment 8: Sensitivity of Model Predictions on μ Prior 326

List of Tables

2.1 The Number of Possible Structures, Interventions and Outcomes For 1–10 Binary Variables . . . 53

3.1 Three Most Frequent Judgment Errors in Experiments 1, 2a and 2b. . . . 85
3.2 Comparison of the Proportion of Interventions of Different Types Selected by Participants and Simulated Learners. . . . 90
3.3 Model Fitting Results. . . . 108

4.1 Percentage of Time Participants Chose Each Intervention . . . 127
4.2 Experiment 3: Regression analyses . . . 128

5.1 Proportion of Edges Correctly Identified by Device Type in Experiments 1, 2, 4 and 5 . . . 165
5.2 Edit Distance Between Consecutive Judgments in Experiments 1, 2, 4 and 5. . . . 165
5.3 Belief Model Fits. . . . 181
5.4 Intervention Model Fits . . . 185

6.1 Experiment 6: Order and Delay Models Compared to Participants' Judgments . . . 223
6.2 Experiment 7: Possible Temporal Order Patterns . . . 228
6.3 Experiment 7: Evidence Sets (1^{st} - 4^{th} Piece of Evidence) for the 8 Devices 228
6.4 Experiment 7: Likelihood Judgment Model Fits . . . 231
6.5 Experiment 8: Main Effects and Planned Comparisons for First, Second and Third Responses . . . 242

7.1 Model comparison . . . 274

8.1 Experiment design . . . 286

A.1 Belief Model Recovery Analysis . . . 320

B.1 Experiment 8: Correlation Between Heuristic Measures . . . 327

Abbreviations

CBN	A **C**ausal **B**ayesian **N**etwork
Ug	**U**tility **G**ain
Pg	**P**robability **G**ain
Ig	**I**nformation **G**ain
Eug	**E**xpected **U**tility **G**ain
CF(S/G/U)	**C**onservative **F**orgetful (**S**cholar/**G**ambler/**U**tilitarian) (Chapter 3)
NS	**N**eurath's **S**hip (Chapter 5)
SE	**S**imple **E**ndorser (Chapters 3 and 5)
WSLS	**W**in-**S**tay, **L**ose-**S**ample (Chapter 5)
Order$_S$	**S**imultaneous **O**rder (Chapter 6)
Order$_N$	**N**onsimultaneous **O**rder (Chapter 6)
Order$_{IV}$	**I**ndi**v**idual **O**rder (Chapter 6)
Delay$_P$	**P**ooled **D**elay (Chapter 6)
Delay$_I$	**I**ndependent **D**elay (Chapter 6)
APD	**A**pproximate **P**airwise **D**ifference model (Chapter 6)
PD	**P**redictive **D**ivergence (Chapter 8)

Symbols

m	A causal model
$\mathcal{M} = \{m_1 \dots m_N\}$	A set of possible causal models
M	The unknown true causal model as a random variable
$\mathbf{X} = \{X_1 \dots X_N\}$	The variables in a causal Bayesian network (Chapters 1–6)
$X_i \to X_j$	A causal connection from X_i to X_j
$X_i \leftarrow X_j$	A causal connection from X_j to X_i
$X_i \nleftrightarrow X_j$	No connection between X_i and X_j
E_{ij}	Unknown edge in a causal model $\in \{1 : i \to j,\ 0 : i \nleftrightarrow j,\ -1 : i \leftarrow j\}$
$\mathrm{pa}_m(X_i)$	The parents (direct causes) of X_i in causal model m
$\mathrm{de}_m(X_i)$	The descendants (direct and indirect effects) of X_i in causal model m
$\mathbf{d} = \{X_1 = \mathrm{x}_1 \dots X_N = \mathrm{x}_N\}$	Data from a single observation/intervention
t	Test or time indices
T	The cardinality of tests or total time in a trial
$\mathcal{D}^t = \{\mathbf{d}^1 \dots \mathbf{d}^t\}$	The data from tests or trials 1 to t
$\mathcal{D}_{\mathbf{c}}$	The set of all possible data that might arise from intervention $\mathbf{c}$
$\mathrm{Do}[X_i = \mathrm{x}]$	"Do" operator (Pearl, 2000). Sets variable(s) X_i to state(s) x
$\mathbf{c} = \{c_1 = \mathrm{c} \dots c_N = \mathrm{c}_N\}$	An intervention setting a subset of $\mathbf{X}$
$\mathcal{C} = \{\mathbf{c}^1 \dots \mathbf{c}^t\}$	Interventions performed on tests or trials 1 to t
$\mathcal{D}_r^t$	"Recent" data (Chapter 5)
$\mathcal{C}_r^t$	"Recent" interventions (Chapter 5)
b^t	The learner's current model belief
$\mathbf{w}$	The parameters of a causal model
w_S	Strength parameter of a CBN

w_B	Base rate parameter of a CBN
α	Softmax parameter for judgments (Chapter 3)
β	Softmax parameter for interventions (Chapter 3)
γ	Forgetting parameter (Chapters 3 and 4)
η	Conservatism parameter (Chapter 3)
θ	Heuristic intervention selection parameter (Chapter 3)
σ	Heuristic belief selection parameter (Chapter 3)
ρ	Heuristic belief inertia parameter (Chapter 3)
κ	Heuristic disambiguation step parameter (Chapter 3)
λ	Search length parameter (Chapter 5)
k	Chosen search length (Chapter 5)
ω	Search behaviour parameter (Chapter 5)
R	Transition matrix for all possible search paths (Chapter 5)
η	Local intervention choice softmax parameter (Chapter 5)
ρ	Focus choice softmax parameter (Chapter 5)
θ	Locale choice softmax parameter (Chapter 5)
ϵ	Noise parameter (Chapter 5)
$A, B, C, \ldots$	Labels for components in dynamic causal systems (Chapters 6–8)
α	Shape parameter for causal delays (Chapter 6)
μ	Mean parameter for causal delays (Chapter 6)
λ	Rate parameter for real-time spontaneous activation (Chapter 6)
t_X	Activation time of component X (Chapter 6)
t_{XY}	The time between activation of components X and Y (Chapter 6)
$t_{X \to Y}$	The casual delay of X causing Y (Chapter 6)
τ	The end of a period of continuous observation/interaction (Chapter 7)
$\mathbf{z} \in \mathbf{Z}_{\mathbf{d}}^{m}$	Possible paths of actual causation (Chapter 7)
$\phi \in \Phi_{\mathbf{z}}^{\mathbf{d}}$	Set of expected events that failed to occur (Chapters 7)
$d_X^{(i)}$	The ith component activation, located at component X (Chapters 7)
$c_X^{(j)}$	The jth intervention, located at component X (Chapters 7)

$b^{(i)}$	The belief upon seeing the ith activation (Chapters 7)
$\mathbf{x}$	Observed object trajectories (Chapter 8)
$\mathbf{s}_m$	Simulated object trajectories under possible world m (Chapter 8)
ϵ	Perceptual noise (Chapter 8)
$\mathbb{E}[f]_x$	Expectation of f with respect to x
$I[f]$	Indicator function, equals 1 wherever f is true

To my brother Douglas

The chapters in this thesis are based on the following articles:

Chapter 3

Bramley, N. R., Lagnado, D. A. & Speekenbrink, M. (2015). Conservative forgetful scholars: How people learn causal structure through interventions. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *41*(3), 708–731.

Chapter 4

McCormack, T., Bramley, N. R., Frosch, C., Patrick, F. & Lagnado, D. A. (2016). Children's causal structure learning. *Journal of Experimental Child Psychology*, *141*, 1–22.

Chapter 5

Bramley, N. R., Dayan, P., Griffiths, T. L. & Lagnado, D. A. (2017). Formalizing Neurath's ship: Approximate algorithms for online causal learning. *Psychological Review*, to appear.

Bramley, N. R., Dayan, P. & Lagnado, D. A. (2015). Staying afloat on Neurath's boat: Heuristics for sequential causal learning. In D. C. Noelle et al. (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, (pp. 262–267). Austin, TX: Cognitive Science Society.

Chapter 6

Bramley, N. R., Gerstenberg, T., Mayrhofer, R. & Lagnado, D. A. (submitted). The role of time in causal learning.

Bramley, N. R., Gerstenberg, T. & Lagnado, D. A. (2014). The order of things: Inferring causal structure from temporal patterns. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 236–242). Austin, TX: Cognitive Science Society.

Chapter 7

Bramley, N. R., Mayrhofer, R., Gerstenberg, T. & Lagnado, D. A. (submitted). Causal learning from interventions and dynamics in continuous time.

Chapter 8

Bramley, N. R., Gerstenberg, T. & Tenenbaum, J. B. (2016). Natural Science: Active learning in dynamic physical microworlds. In A. Papafragou, D. Grodner, D. Mirman

& J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 2567–2573). Austin, TX: Cognitive Science Society.

Chapter 1

Introduction

> "*All reasonings concerning matter of fact seem to be founded on the relation of cause and effect. By means of that relation alone we can go beyond the evidence of our memory and senses.*"
>
> — DAVID HUME

The story of model-based cognition begins with the discovery of causality. Born into a "blooming, buzzing confusion" (James, 1890, p462), we discover our first causal relationship in our ability to affect this sensory barrage. Opening our eyes brings light. Crying brings food. Motor actions affect sensations. A control loop of action and effect is established that drives subsequent learning. As we learn, we generalise from our own ability to affect things, populating our model of the world with multiple entities and causal relationships. By adulthood, this control loop has produced a rich model of a causal world filled with objects and forces, relata and relationships, with ourselves as the original and most important player. This is a "just so" story (Kipling, 1902), but it illustrates an intuitive idea. Causality and action are closely related, together providing the building blocks and tools needed to build a rich and productive ontology. Causal models are valuable to us because they let us venture, virtually, away from the here-and-now while maintaining a relationship with reality. We can travel forwards — predicting what will happen in the upcoming seconds, days and years; backwards — generating explanations for encountered phenomena; and sideways — imagining how things would play out if circumstances were different (Woodward, 2003). By simulating potential actions before taking them, causal knowledge lets us act flexibly and intelligently, choosing actions likely to get us where we want to go first try, even as our circumstances and goals shift. This metaphysical mobility is so embedded in higher-level cognition it is

easy to take for granted, but its success as a cognitive strategy depends on the match between the model in the head and the causality in the world. Therefore, in this thesis, I explore two closely related questions: (1) How do people learn causal models through their interactions with the world? And: (2) How do people interact with the world in order to learn about its causal structure?

I approach both problems from the perspective of rational analysis (Anderson, 1990). This means I work from the top down; first laying out a computational level perspective on causal structure induction and interventional learning (Marr, 1982), using this alongside behavioural data to develop models of psychological representations and processes. Learning a causal model likely to have produced the evidence of one's senses is a classic reverse-engineering problem. Like many of the problems faced by cognition, it is one of induction — any causal hypothesis could be falsified by new evidence (Hume, 1740) — and fundamentally ill-posed — the data massively under-determines the correct solution. As we will see in Chapter 2, this means that normative inference can be understood through the mathematics of subjective probability theory and Bayesian inference over a large hypothesis space of possible structural models. I will show that learning a causal model is harder than learning a descriptive or associative one, largely possible only because we are active participants in the world we seek to understand. Choosing how to intervene on the world to learn its structure is a problem of active learning where the goal is to choose actions expected to reduce one's uncertainty (Settles, 2012). Thus, I will analyse people's active causal learning from an information theoretic perspective, where normative active learning can be defined as taking actions expected to reduce uncertainty about the true causal model as quickly as possible (Miller, 1984; Shannon, 1951).

In general, rational analysis will serve to reveal the prohibitive complexity of optimal causal learning and action selection in the real world. Thus, throughout the thesis, I will develop a "resource rational" (Griffiths, Lieder, & Goodman, 2015; Simon, 1982) perspective on how people refine their causal models and intervene effectively on the world despite their processing and representational limitations.

In addition to being under-determined by the evidence, the brain is just a small part of the world it seeks to understand. This puts hard constraints on causal representation in terms of storage and efficiency. People's causal models must be useful for making causal inferences under uncertainty in real time, meaning that representations must provide a compact and efficient abstraction that keeps what is predictively useful while discarding the rest. As we will see, one way to do this is by learning a probabilistic

map of causal relationships between variables that can accommodate noisy and imperfect relationships, and that we can model using a graphical model formalism called a *causal Bayesian network*. Another aspect is hierarchical structure, through which cognisers capture commonalities between domains that bootstrap inference (Griffiths & Tenenbaum, 2009; Tenenbaum, Kemp, Griffiths, & Goodman, 2011), allowing for strong causal judgments in familiar domains after even a single data point.However, causal knowledge also extends beyond contingency relationships between variables. We experience the world as extended in space and time. So to attribute causes to effects and to act effectively, we must relate our abstract beliefs to these dimensions. Thus, in the later chapters of the thesis I will explore the role of time and mechanism knowledge in learning and representation.

Over the thesis as a whole, a consistent perspective will emerge. Human causal learning is a stochastic and gradual process, shaped as much by people's current causal model when considering alternatives as by the ground truth they hope to approach. In particular, in Chapter 5, I develop a scalable model of incremental structure learning, named after the Neurath's ship metaphor in philosophy of science (Neurath, 1932; Quine, 1969). The metaphor relates the challenges of theory change to those of fixing a ship while at sea, and the model embodies this process by showing how a learner can gradually refine a global structural model through small incremental changes (as they think and gather evidence). The model has a close formal relationship with sample-based algorithms for approximation in machine learning, specifically particle filters and Markov Chain Monte Carlo. Thus, it captures the idea of *resource rationality* (Griffiths et al., 2015; Simon, 1982), showing how learners trade accuracy against computational cost in ways that allows them to learn as well as possible even in highly complex domains. I show that *Neurath's ship* outperforms a number of alternative proposals in the literature in describing behavioural patterns in discrete-trial interventional causal learning. In Chapters 6 and 7, I show that we see similar patterns of incremental construction in learning continuous time settings. In the General Discussion (Chapter 9), I will propose that this approach is not limited to learning structure, but that variants of this approach can explain human success at learning complex representations in general.

In parallel to the insights into learning and representation, the thesis will develop a novel perspective on interventional active learning, as a process of optimal *self-teaching* rather than optimal *information gathering*. From this perspective, the learner's goal is to find actions that optimally support their own bounded and limited ability to learn, rather than maximising information per se. In Chapter 5, we will see that people choose

interventions in complex causal systems to learn about manageable subspaces of the problem, such as a single relationship at a time, the effects of a single variable, or whether their current favoured hypothesis is correct. When intervening on causal systems in time (Chapter 7), we will see that people space out their interventions in a way that structures their learning experience as a whole, creating approximately independent "trials". And, learning by intervention in simulated causal worlds, we will find that people use their actions creatively to construct informal "experiments", that test properties of interest one at a time while minimising "causal noise". The local uncertainty schema I develop captures formally why human learners behave in the ways they do, explaining a range of common behaviours in the active learning literature that appear inefficient on first analysis, including preference for confirmatory testing; and repetition (Schwartzman, 2012) and failure to control for confounds (Chen & Klahr, 1999).

More broadly, this thesis will contribute to open questions in philosophy, computer science and psychology. In philosophy, my *Neurath's ship* model of incremental causal theory change puts formal flesh on anti-foundationalist ideas about the progression of scientific theories (Kuhn, 1962; Lakatos, 1976; Thagard, 1992) providing a plausible picture of how complex theories grow and change through the combination of evidence and thought. In computer science, the data I present provide rich evidence about how humans solve the problems associated with learning generative causal models of the world. This has potential to guide the development of human-like artificial intelligence. In psychology, this thesis contributes to debates about Bayesian rationality, heuristics and approximation. My models of psychological processes underlying causal learning not only explain how humans can learn in ways that respect uncertainty yet scale up to problems of real-world complexity (Griffiths et al., 2015; van Rooij, Wright, Kwisthout, & Wareham, 2014), but also predict common behavioural phenomena including probability matching, anchoring and confirmation bias.

The structure of this thesis is as follows: Chapter 2 develops the computational level picture of causal inference and intervention selection, introducing rational analysis, probabilistic graphical models, exact and approximate Bayesian inference and active learning. Chapter 3 explores interventional causal learning from contingency information in experiments and modelling. It compares three measures of intervention choice, and compares an idealised Bayesian learning model to bounded variants, finding that learners are both conservative in their judgments and forgetful about evidence relative to Bayesian norms. At the start of Chapter 3, I provide mathematical details of the CBN formalism in a text box for reference in reading this and subsequent chapters. Chapter 4 explores the

developmental trajectory of interventional causal learning in 5- to 8-year-olds. It shows that all but the youngest children choose interventions that are more useful than chance and that the quality of their intervention choices predict judgment accuracy. But, as with adults, judgments are indicative of forgetting and neglect of older evidence and intervention choices are repetitive and confirmation-biased.

Chapter 5 explores how people scale up to learn in harder problems with complex causal structures and substantial noise. In this chapter I develop and test my *Neurath's ship* process model of bounded structure inference and locally focused intervention choice, showing that it describes participants' behaviour better than a number of alternatives. Chapter 6 focuses on the role of time in interventional structure learning, finding that participants make systematic use of event order and inter-event intervals, and that we can again describe their behaviour as based on entertaining and adapting a single candidate structural hypothesis. Chapter 7 explores how people make interventions in dynamic causal systems in real time, showing that they are good at using interventions to break a continuous learning period into separate approximate "trials", and make structure judgments that, again, reflect the incremental construction of a single causal model hypothesis. Chapter 8 looks at active learning in rich simulated physics worlds with the goal of identifying their latent causally relevant properties of objects like their masses and forces of attraction and repulsion. I show that learners use their actions to construct "natural experiments" that use the laws of physics to strongly reveal target properties of objects while minimising confounding information. In Chapter 9, I bring all these results together and construct a cohesive picture of active causal learning and representation in cognition, and comment on the implications of this work in philosophy, computer science and psychology.

Chapter 2

Causal cognition

> "*Our grasp of the world — the way we mirror its causal structure — is at the mercy of the inferential tools we have in the brain.*"
>
> — JAKOB HOHWY

Studying active causal learning is difficult because it is so deeply embedded in cognition, cross-cutting core cognitive competencies such as perception, probabilistic and approximate inference, abduction and abstraction. Indeed, this thesis is itself an exercise in active causal learning: it will involve constructing models of how people construct causal models, and testing people to see how they test the world. Similar to how I will argue people learn causal structure, I will build on rich prior knowledge (existing research), gather new evidence, and integrate the two via Bayesian statistics with the goal of refining a picture of human active causal cognition. The complexity of the topic means it is important to be clear about the methods of analyses, and to delineate the theoretical and empirical questions. Accordingly, in this first chapter I introduce necessary apparatus for studying causality and active learning, and survey the relevant literature. In particular, I introduce computationalism, rational analysis, probabilistic graphical models, exact and approximate Bayesian inference and active learning. Using these tools, I then set up the problem of interacting with the world and learning a model of its causal structure.

2.1 Computationalism

The 20^{th} century saw the invention of the computer, and along with it the rise of computationalism in psychology. Computationalism is the thesis that the mind/brain is

fundamentally an information processing device, meaning that psychological processes like learning, thinking and decision making are explicable in computational terms. The success of this approach is evident in the rapid co-evolution of the cognitive and computer sciences in the last 50 years. Many algorithms in contemporary machine learning and artificial intelligence — e.g. image and voice recognition systems, self driving cars, recommendation engines — make use of insights about how analogous computations are achieved in the brain. In parallel, as computational power and know-how has increased, so too has the strength of the computationalist perspective; as better cognitive models have enriched our understanding of cognition.

One reason for the success of computationalism in psychology is the clarity it lends to the formation and refinement of theories. In studying vision, Marr (1982) described three levels at which one can describe information processing systems: (1) At the *computational level*, explanation focuses on the abstract problem being solved by a system. What is the system designed or evolved to do? What inputs should it map to what outputs? (2) At the *process level*, explanation focuses on how a system solves the problem. What does it do with the inputs to come up with its outputs? What representations and algorithms are involved? (3) Finally, at the *implementational level*, explanation focuses on how a system is physically implemented. What is it made from: gears and levers; electrical circuits; neurons? How are the parts arranged and connected such that they realise the representations and algorithms?

Marr's scheme has since become central to computational analysis. It captures how understanding of cognition on one level informs and constrains inference about the other levels. Only certain algorithms can solve a given computational problem, and only certain physical architectures are appropriate for realising a given algorithm. As cognitive scientists, we are interested in understanding the problems solved by cognitive agents (the computational level), but even more interested in how they solve them (the process level). One project is to build up our understanding of cognitive processing via close analysis of neuroanatomical implementation. However, we often want to keep our participants alive, and the complexity of the brain means that higher level cognition is still largely beyond the reach of this bottom up explanation. Thus, a parallel project is to work from the top down, first refining understanding of the problems the brain solves, then theorizing about the representations and algorithms it uses to solve them. In this thesis I take this top-down approach.

2.2 Rational analysis

Related to Marr, Anderson (1990) laid out a method for rational analysis of cognition that has been widely adopted in cognitive science. It rests on the assumption that the pressures of evolution and development have rendered cognition close to optimal given its constraints. Anderson proposes we should start with a first-principles description of the problem being solved by the agent at the computational level, in terms of its goals and environment, making minimal assumptions about its computational limitations. We work out the optimal solution to this problem, then compare this to participants' behaviour in experiments. Where there are discrepancies, we use these to refine our understanding of the problem being solved either by: (1) incorporating more plausible characterisation of the computational constraints; (2) reconsidering the agents' goals, or (3) refining our model of the environment. By iterating this process, Anderson argues we can get closer to understanding cognitive processing. Building on these ideas and Marr's hierarchy, Chater and Oaksford (1999) emphasise that there are often a number of "deep local optima" — i.e. multiple plausible algorithms that approximately solve a refined computational problem. This means that interpreting deviations from model predictions involves striking a balance between refining one's idealisation of the problem being solved, and considering candidate process level models that can explain deviations from these optima. As we will see, optimal inference in many cognitive domains, not least causal inference and active learning, is radically intractable, requiring practically infinite storage and processing power (Brighton & Gigerenzer, 2012). This means there can be a large gulf between rational models and what could plausibly be implemented by the brain. Thus, in recent years, there has been an increased emphasis on consideration of the representational and algorithmic implications of ostensibly "rational" models (Jones & Love, 2011) and focus on identifying rational approximations that can account for human successes without positing inhuman computations (Griffiths et al., 2015).

2.3 Bayesian inference

In the last few decades, rational analysis has shed light on the character of the computational problems faced by cognition. Where early computationalist ideas about mind treated it as fundamentally engaged in logical symbol manipulation (McCulloch & Pitts, 1943; Newell & Simon, 1972; Whitehead & Russell, 1912), there has since been a "probabilistic turn" (Chater, Tenenbaum, & Yuille, 2006). The anecdote goes that in the

1960's, MIT's Marvin Minksy assigned a graduate student the problem of solving computer vision as a summer project (Blackmore, 2000). By the end of the summer, all they had established was that the problem was much harder than initially thought. Everyday problems faced by cognitive agents — like recognising objects, and understanding natural language, and motor control — long resisted the grasp of what, following Haugeland (1989), is now often termed "good old fashioned AI". The world is complex, noisy and under-constrained by the evidence of the senses. This means it is rational to be uncertain, and to respect this uncertainty in one's representations and inferences (Chater & Oaksford, 1993; Oaksford & Chater, 2007). In general the inferences required of cognition have turned out not to be *deductive*, but rather *inductive* in the sense that beliefs about the world are always uncertain and falsifiable (Goodman, 1955; Hume, 1740), and *abductive* in the sense that a central goal is to learn parsimonious generative models that can explain lots of encountered evidence (Peirce, 1955). The problem of induction, and to some extent abduction, have turned out to be problems of probabilistic inference, addressable via probability theory (Carnap, 1962).[1]

Probability theory allows us to treat uncertain quantities as random variables, meaning that while we do not know their true value, we can maintain a probability distribution over their possible values that can change as we gather more evidence. Bayes' theorem (Bayes & Price, 1763) is a consequence of the axioms of probability and provides a calculus for rational inference under uncertainty. From a Bayesian perspective, learning is the process of updating one's subjective probabilistic beliefs about the true state of some part of the world, where the ground truth is treated as a random variable X. Its possible values $x \in X$ cover all possible hypotheses about about how things might be. A Bayesian learner, having observed evidence d, updates their prior probability distribution $P(X)$ into a posterior distribution $P(X|d)$, by multiplying prior $P(X)$ and likelihood $P(d|X)$ and normalising by the average likelihood of the data across all the possible values of x:

$$P(X|d) = \frac{P(d|X)P(X)}{\sum_{x_i \in X} P(d|x_i)P(x_i)}. \tag{2.1}$$

The posterior from one learning instance becomes the prior for the next (e.g. $P_t(X) = P_{t-1}(X|d_{t-1})$), and this process continues as more evidence is received.

[1] Although probability theory does not solve the strongest forms of the philosophical problem of induction. Probabilistic inference still rests on the assumption that the future will resemble the past, an assumption that itself can only be justified inductively. Thus, the attempt to justify induction is circular.

We can motivate this idea with a simple example. What is the prior probability that a randomly chosen person plays in a rock band? Suppose you think it is quite low, say $P(x_{\text{band}}) = \frac{1}{200}$. Suppose you now learn that this person likes rock music $d =$"likes rock music". The chance that someone who is actually in a rock band likes rock music is probably very high, say $P(d|x_{\text{band}}) = \frac{99}{100}$. But then liking rock music is not uncommon. Suppose among those who don't play in rock bands, the chance of liking rock music is still around $P(d|x_{\text{not}}) = \frac{1}{5}$. Now we can use Bayes' theorem to update our belief that the person in question is in a rock band, given that we now know their music preferences. Bayes' theorem tells us that we should rationally increase our belief that this person is in a rock band from $\frac{1}{200}$ to around $\frac{1}{40}$. With sufficient evidence, an idealised Bayesian learner's subjective beliefs eventually approximate the ground truth, in this case with $P(x_{\text{band}})$ approaching 1 if the person actually is in a rock band and 0 if they are not.[2]

In general, the problem of updating one's probabilistic beliefs about the world is what is known as an "inverse problem" (Ambartsumian, 1929; Tenenbaum et al., 2011). This means one must work backward from evidence to infer the probability of the causal factors that produced it. Inference about the state of a single variable is a fairly minimal example of this. However, once we begin to consider Bayesian inference about the causal model relating multiple variables, our "X" of interest covers an increasingly large space of possible generative models. I depict this in Figure 2.1, where causality flows both around, in the world, and outwards from the world to the data it produces. In contrast, inference inverts this process, flowing "inward" from the data to refine one's internal causal beliefs (which, in turn, contain virtual causality, hopefully mirroring the causality in the world).[3]

In order to understand how people learn a causal model and use it to guide action, we must think about the form that their causal models take. It is clear that the goal of causal learning is to come up with a model of the world that makes it predictable and, ultimately, exploitable. I depict this idea in Figure 2.1. We can affect the world with our actions (hand symbol) and want to do so in ways that reap rewards (find food, minimise pain etc). However, we need to know something about the world in order to choose actions likely to bring rewards. Some of this is "model free", in the sense that reinforcing actions that were rewarding in the past is often a good start. Indeed, a large portion of the human and animal learning literature focuses on model

[2]Provided, of course, that the hypotheses are distinguishable and the hypothesis space contains the ground truth.

[3]While I depict this as a single causal model, from a Bayesian perspective this should be a probability distribution over all possible causal models.

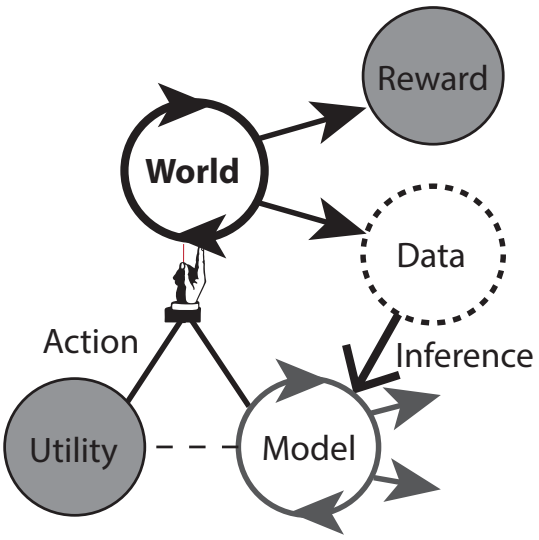


FIGURE 2.1: Schematic of active causal learning. The *world* produces *data* and *reward*. *Data* are used to refine a *model* of the world. One's *model* and *utilities* drive *action* that affects the *world*, producing future *data* and *rewards*. Note that I assume utilities are indirectly related to *rewards* via the *model* (dashed line).

free forms of learning capturing many behaviours as direct responses to stimuli (e.g. Mackintosh, 1983; Skinner, 1990; Sutton & Barto, 1998). Unfortunately, one can only get so far without a model (Daw, Gershman, Seymour, Dayan, & Dolan, 2011). When rewards move or deplete, or actively try to outmanoeuvre you, reinforced behaviours become useless or maladaptive. Furthermore, it is hard to know what will be valuable in the future, so reward dependent learning can easily fail to equip one for the future. Intuitively then, a causal model provides flexibility, enabling the bearer to navigate the world successfully wherever they end up wanting to go. This flexibility depends on learning a representation that mirrors the way the world actually works (Hohwy, 2013); one that — setting aside metaphysical claims about causation — captures the *true* causal relationships. However, the model must also be compact both because it must fit inside a brain much smaller than the world it imitates, and because it must support online inference about the current state of the world.

2.4 Probabilistic graphical models

Before considering causal representations, it is helpful to consider how we can represent relationships between variables in general. As we have seen, probabilistic inference is straightforward to demonstrate in toy problems. However, even my simple example above presupposed knowledge about the relationships between the variable of interest

and the evidence (e.g. between liking rock music and being in a rock band). In general, a probabilistic brain needs a representation that supports these kinds of *conditional probability* judgments relating arbitrary variables, so that learning about one thing can have the right knock-on implications on the probabilities of whatever else it is related to.

A naïve way to do probabilistic inference about lots of variables is to keep track of the probabilities of all the combinations of all the states of all the variables, averaging over all the unobserved variables when making inferences. Unfortunately, as the space we are interested in doing inference over gets more complex, the amount of work and storage needed to track and calculate probabilities increases much faster, rapidly becoming infeasible. Even if, implausibly, one's entire understanding of the world could be captured by 100 binary variables, one would would need to keep track of 2^{100} (which is about 10^{30}) probabilities to make basic inferences. Figure 2.2b illustrates this idea with just 10 variables. Worse, when using one's knowledge to make inferences, one would often have to average over very large numbers of possibilities. For instance, suppose the two variables from the toy example above feature in our 100 variable representation. Concretely, let X_1 be "a person likes rock music" and X_2 be "a person is in a rock band". To get the marginal (i.e. average) probability of someone being in a rock band, one would still have to sum over all combinations of the other 99 variables $P(X_2) = \sum_{X_1} \sum_{X_3} \dots \sum_{X_{100}} P(X_1, X_3 \dots X_{100})$. To get the updated probability that a person is in a rock band having learned that they like music, one would have to condition on $X_1 = 1$ and marginalise over all combinations of the other 98 variables $P(X_1|X_2 = 1) = \sum_{X_3} \dots \sum_{X_{100}}, P(X_2, \dots, X_{100}|X_1 = 1)$, and so on. Clearly this is not feasible for complex beliefs involving many variables that might take many or a continuum of states. This combinatorial explosion is a serious and pervasive problem for probabilistic inference. Luckily, this complexity can be managed through finding *structure* that simplifies the problem, representing only those relationships that are essential for retaining the probabilistic information (Bishop, 2006).

Computer scientists model structure in large probability distributions using probabilistic graphical models. Graphical models encode the relationships between variables in a lightweight way by storing probabilities only between those variables that are directly related. If there are still too many relationships to store, graphical models can trade off veracity for efficiency by only representing some of the direct relationships (e.g. the strongest ones). In general, the less densely connected the variables in the graphical model, the easier it is to make probabilistic inferences (Barber, 2012).

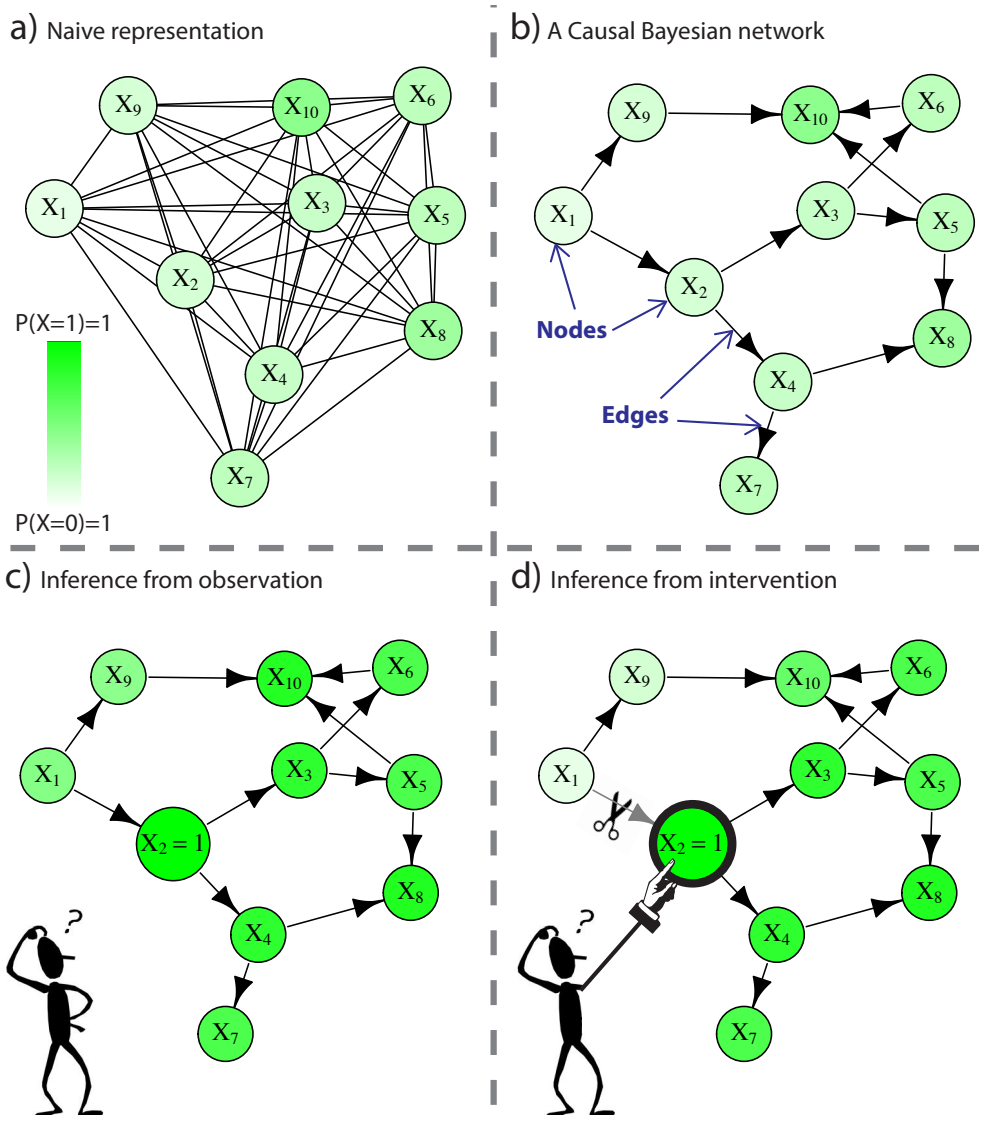


FIGURE 2.2: Graphical model examples. a) Visualisation of a joint distribution over 10 variables. Node shading denotes marginal probability that each variable is "on" (i.e. takes the value 1). Lines indicate naïve assumption that every variable depends on every other variable. b) A (causal) Bayesian network for the same 10 variables. c) The marginal probability of all unobserved variables X_1, X_3–X_{10} after observing X_2 being "on" (taking the value 1). In this example, all variables have a small *base-rate* probability of being "on" and causal relationships have a positive *strength* (e.g. they raise the probability of their effects). Children and parents of X_2 are more likely to be on too. Note X_2 node is enlarged for readability. d) An example of the marginal probability of all unobserved variables after an intervention that turns X_2 "on". Only children of X_2 are more likely to turn "on". Grey arrow and scissors indicates that the intervention overrides the normal connection from X_1 to X_2.

A common class of probabilistic graphical model is the Bayesian network. Bayesian networks capture probabilistic structure in terms of conditional probabilities and represent these using an acyclic graph. Variables are represented by nodes and conditional probabilities are represented by arrows (called "edges", see Figure 2.2b). For a given

node, other nodes that can be reached by travelling forward along edges are its "descendants", and any node that can be reached by travelling backward along edges are its "ancestors". Immediate descendants are called "children" and immediate ancestors are called "parents". Bayesian networks are defined by the Markov condition, which states that each node is independent of all of its non-descendants given its parents. Crucially, this reduces the number of variables that must be considered when doing inference.

We can illustrate this by introducing a few more binary variables to our example. Suppose X_3 is "a person has tinnitus" and X_4 "a person spends a lot of time travelling". Suppose that X_2 (a person is in a band) is probabilistically related to both X_3 and X_4. That is, people in rock bands are both more likely to get tinnitus (e.g. because they are exposed to lots of noise) and to spend a lot of time travelling (e.g. on the road between gigs).[4] Suppose also, that there is no direct relationship between X_3 and X_4 — e.g. between whether someone has tinnitus and how much time they spend travelling, once we have accounted for any effects due to playing in a band. This means X_3 is "conditionally independent" of X_4 given X_2 (we can write this as $X_3 \perp\!\!\!\perp X_4|X_2$). This means that we can construct a Bayesian network relating these variables by replacing the joint distribution over all 8 combinations of $P(X_2, X_3, X_4)$ with a product of simpler conditional distributions that omit the (non)relationship between X_3 and X_4. However, there are actually several ways of doing this. We can define the probabilities of X_3 and X_4 as conditional on X_2, giving $P(X_2)P(X_3|X_2)P(X_4|X_2)$ corresponding to a Forking structure $X_3 \leftarrow X_2 \rightarrow X_4$. Alternatively, we could start from X_3 defining X_2 conditional on X_3 but keeping X_4 conditional on X_3, giving $P(X_2)P(X_3|X_2)P(X_4|X_3)$ corresponding to a Chain $X_2 \rightarrow X_3 \rightarrow X_4$ structure. Finally, we could go the other way, defining X_2 conditional on X_4, giving $P(X_4)P(X_3|X_4)P(X_2|X_3)$, also corresponding to a Chain structure but running in the opposite direction $X_2 \leftarrow X_3 \leftarrow X_4$. Formally, these alternatives are known as *Markov equivalent* (Glymour, 2001; Pearl, 1988).

Figure 2.3 gives a worked example of this idea. 2.3a provides some fictional data: frequencies of all combinations of the X_2, X_3 and $X_4 \in$ [0=absent, 1=present] for 10,000 people. 2.3b gives a measure of the degree of probabilistic dependence between the variables, based on these data.[5] This reveals that all three variables are (unconditionally) dependent but that, when conditioning on X_2, X_3 becomes conditionally independent

[4] In general the relationships need not be positive, all that matters is that the probability distribution for one variable differs depending on the state of the other, for at least one setting all the other variables in the network.

[5] I use *Mutual information* (Cover & Thomas, 1991) to measure dependence here. Mutual information captures how much information one variable provides about another and goes to zero if two variables are independent.

of X_4, capturing the idea that tinnitus is unrelated to travelling after accounting for whether someone plays in a band. Figure 2.3c shows how we can parametrise a causal network from these data, and how this depends on how we choose to break up the distribution. The probability of each variable is a function of its parents, and in this binary case we can estimate these by simply counting the proportion of times variables took each state conditional on their parents' states.[6] Variables with no parents, called "root nodes", are simply defined by their marginal probability. Focusing on Figure 2.3c i., X_2 has no parents, so we set it to 0.2 reflecting that, in the data, it takes the value 1 20% of the time. X_3 and X_4 both have X_2 as parent. Thus, they get separate values for cases where X_2 is 0 or 1. If a variable has more than one parent in the graph, we do this for all combinations of the values of the parents. The result is a parametrised Bayesian network, where every variable has a probability given the state of its parent(s). With these set up, the network can be used to reason forward from parents to children (by simply reading off the conditional probabilities) or backwards (by using Bayes rule) from children to parents. For larger networks there are a range of efficient algorithms for performing inference (Barber, 2012). These typically have a computational cost that scales with the density of the connections between the nodes rather than the total number of nodes.[7] It is important to note that even though different possible structures have different root nodes, and different links with different strengths running in different directions (e.g. Figure 2.3c ii. and iii.), they all have the exact same dependency structure, licensing the same marginal and conditional inferences.

Figure 2.4 shows the various Markov equivalence classes for three variable networks. In general, Bayesian and statistical techniques for learning causal Bayesian networks are indifferent about the direction of some of the edges in a model. Rather they identify a Markov equivalence class of possible models, with the orientation of some edges chosen based on prior knowledge about plausible causal direction, or set at random. The Bayesian approach is to treat the true structure as a random variable and update a probability distribution over all possible structures as data is encountered (Cooper & Herskovits, 1992), the difficulty with this being that the hypothesis space of models tends to be very large. There are also algorithms that learn Bayesian network structure by performing statistical independence tests between pairs of variables conditioning on others (e.g. Heckerman, Geiger, & Chickering, 1995; Spirtes, Glymour, & Scheines,

[6] This bases the estimate on the empirical distribution, which is the best maximum likelihood estimator for Bayesian network conditional probability tables (Barber, 2012).

[7] Technically, they scale with the largest fully-connected sub-graph once the network has been converted to a Markov (undirected) network through *moralisation* (Lauritzen, Dawid, Larsen, & Leimer, 1990), a procedure whereby Collider subgraphs become fully connected.

a) Contingency

X_2	X_3	X_4	Count
0	0	0	71207
1	0	0	1919
0	1	0	714
1	1	0	71
0	0	1	7899
1	0	1	17378
0	1	1	73
1	1	1	739

b) Dependency

$X_2; X_3$	0.00225766
$X_2; X_4$	0.24973914
$X_3; X_4$	0.00376807
$X_2; X_3 \mid X_4$	0.00151891
$X_2; X_4 \mid X_3$	0.00000850
$X_3; X_4 \mid X_2$	0.00000000

c) Three Markov equivalent networks

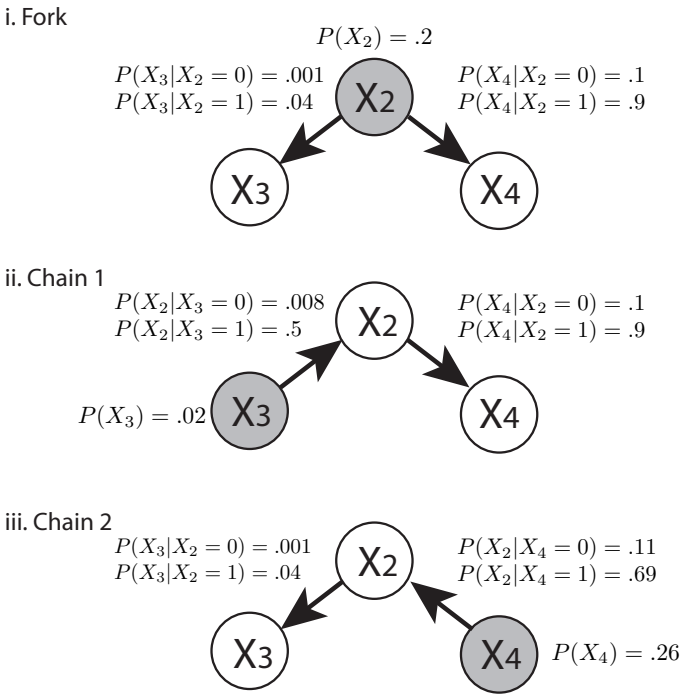


FIGURE 2.3: An example of Markov equivalence. a) Fictional data showing the frequency of different combinations of X_2 (being in a band), X_3 (having tinnitus) and X_4 (spending lots of time travelling). b) The *Mutual information* (Cover & Thomas, 1991) between all pairs of variables. c) Three parametrised Bayesian networks consistent with the data. Each probability is for the variable in question taking the value 1. Grey indicates the root nodes.

1993), using the patterns of dependence to orientate some edges. For example, directionality in Collider sub-graphs can be identified because they imply a unique set of (in)dependencies (Figure 2.4).

2.5 Causal Bayesian networks

Bayesian networks, with their directed connections and close relationship with normative inference under uncertainty, are a promising formalism for modelling people's beliefs about the causal structure of the world. However, the directed edges merely represent *epistemic* relationships and need not run from cause to effect. This means that they do not provide the "guide to life" (Cartwright, 2001, p242) we need from a causal representation. For example, interpreted causally, Figure 2.3c ii. and iii. licence worrying inferences, such as that exposing yourself to loud noises until you develop tinnitus, or travelling a lot, might make you want to join a rock band. In the 80's and 90's, Judea Pearl and others applied the Bayesian network formalism to causal representation by

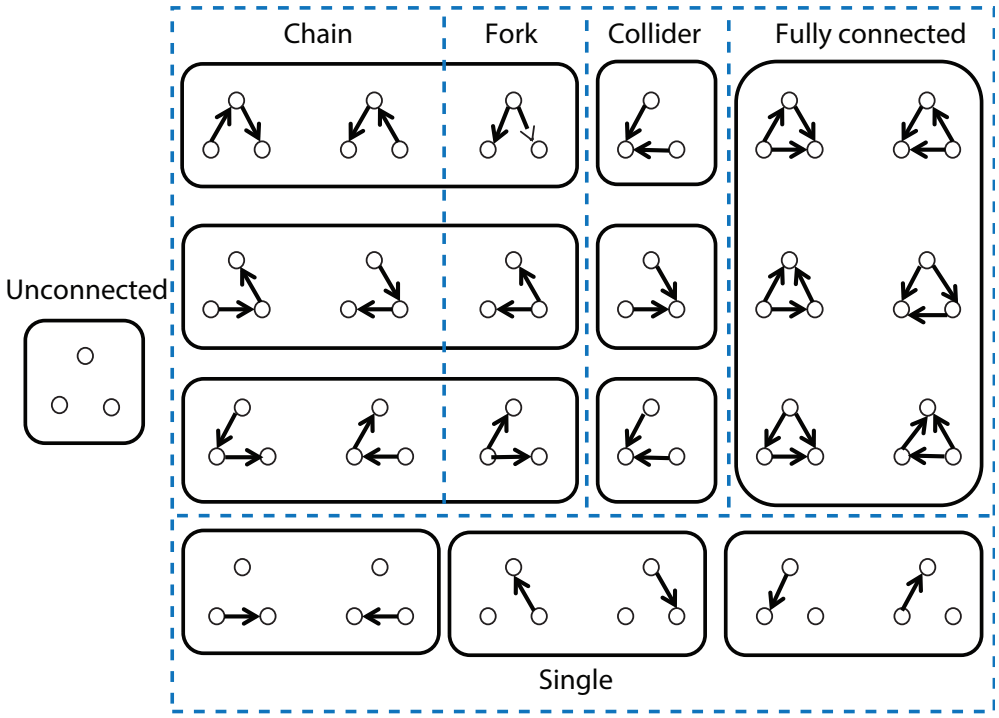


FIGURE 2.4: All 25 directed acyclic graphs on three variables. Black boxes indicate Markov equivalence classes, i.e. types of model that cannot be distinguished from covariational information alone. Blue dashed boxes denote types of structure and the labels I adopt throughout the thesis.

interpreting the directed edges *ontologically*, as probabilistic cause–effect relationships (Pearl, 1988, 2000; Spirtes et al., 1993).[8] Bayesian networks in which the directed edges are interpreted as causal relationships are known as *causal Bayesian networks* (hereafter CBNs). By their nature, the *ancestral order* of the variables in a CBN must correspond to their temporal order. This means the variables that have no parents must be things that come first in time, with their children coming later and so on.

As a core part of the CBN project, Pearl developed an interventional "Do calculus", that formalises the concept of intervening on a causal system (Pearl, 2000). Interventions are situations where one or more of the variables in the model are fixed, exogenously, to one of their possible values. This is a natural way of capturing the idea of reaching into a causal system and manipulating it. Interventions differ from observations in that they affect causally downstream variables, but have no effect on the normal causes of the fixed variable(s). For instance, in our example, if we orientate the edges in the natural causal direction, the model captures what would happen if you intervened, e.g. that joining a a rock band affects someone's probability of developing tinnitus, but that

[8] Pearl actually provides two largely equivalent formalisms. In the first, the relationships themselves are probabilistic. In the second, the relationships are deterministic but indeterminacy is introduced through the inclusion of unknown exogenous influences. I adopt the former formalism but the theory and modelling in this paper is generally compatible with either.

giving someone tinnitus does not affect their probability of joining a rock band. As we will see, this means interventions can be used to break the deadlock between Markov equivalent structures.

CBNs have been widely adopted by psychologists interested in how people learn and reason about causality (e.g. Coenen, Rehder, & Gureckis, 2015; Gopnik et al., 2004; Griffiths & Tenenbaum, 2009; Kemp & Tenenbaum, 2009; Lagnado & Sloman, 2002, 2004, 2006; Lagnado, Waldmann, Hagmayer, & Sloman, 2007; Lee & Holyoak, 2008; Mayrhofer & Waldmann, 2011; Meder, Mayrhofer, & Waldmann, 2014; Oppenheimer, 2004; Oppenheimer, Tenenbaum, & Krynski, 2013; Rehder, 2003; Rehder & Hastie, 2001; Sloman, 2005; Sloman & Lagnado, 2005; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003; Waldmann, Hagmayer, & Blaisdell, 2006). They have also entered into philosophical debate, providing an interventionist perspective on how one might formally ground causal claims (e.g. Danks, 2014; Woodward, 2011). In psychology, the CBN approach has been successful in capturing simple everyday examples of reasoning that seem inherently causal yet are hard to capture with a purely associative model (Holyoak & Cheng, 2011). In particular, the CBN framework captures the ways in which people's judgments exhibit "explaining away" — where multiple potential causes of a common effect compete as explanations. For example, suppose you learn that someone suffers from tinnitus. Other things being equal, this should increase your belief that they play in a rock band. However, if you now learn that they work in an extremely noisy factory, your suspicion that they play in a rock band will be reduced. CBNs also capture another phenomenon called "screening off", in which the information one variable provides about another is reduced to the extent that one already knows the values of any causally intervening variables. This is like saying that, if you already know someone is in a rock band, learning that they also like rock music will have less influence on your expectation that they have tinnitus, removing it completely if you believe the association between liking rock music and having tinnitus is entirely mediated by playing in a rock band.

CBNs have also helped make sense of people's judgments of causal strength from covariation information. Before the development of the CBN there was a long-running debate about how people go from contingency information to judgments of the strength of a causal relationship between binary variables. An early proposal was that causal strength judgments reflect the change in probability of an effect given a cause — known as ΔP (Allan, 1980; Jenkins & Ward, 1965; Lopez, Cobos, Caño, & Shanks, 1998). After finding systematic judgments that ΔP could not capture, Cheng (1997) proposed that causal strength estimates were better understood through the concept of *causal power*,

which captures the probability that a cause brought about its effect, after accounting for the chance it was caused by other, background, factors. Griffiths and Tenenbaum (2005) found that putative strength judgments across a number of experiments could often be better explained by the idea that learners are actually making *structure judgments* within the CBN framework. So, rather than assuming a relationship and judging its strength, participants' estimates reflected the rational judgments of the probability of a causal relationship given the evidence they had been shown. *Causal power* turns out to embody a natural way of parametrising CBNs on binary variables, known as noisy-OR (Pearl, 1988). Noisy-OR considers the probability of an effect to be the probability that at least one of its causes was effective. Subsequent work on causal cognition has frequently assumed both the CBN framework and the noisy-OR parametrisation (e.g. Coenen et al., 2015; Lagnado & Sloman, 2006; Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2008; Yeung & Griffiths, 2015). The example in Figure 2.2b is, in fact, a noisy-OR parametrised causal Bayesian network, in which variables activate by chance with probability 0.1, and active causes have a strength of 0.8. Thus those with many ancestors (e.g. X_{10}) are more likely to be activated either by chance or by a parent, than those with fewer ancestors (e.g. X_1).

Despite its successes, it should be noted that CBN theory also overestimates people's sensitivity to contingency information. A number of papers have shown that people often make systematic deviations from normativity from the perspective of CBN theory (Mayrhofer, Goodman, Waldmann, & Tenenbaum, 2008; Park & Sloman, 2013; Rehder, 2014; Walsh & Sloman, 2008). CBNs are also poor choices for representing certain probabilistic inferences such as inference about mutually exclusive causes (see Fenton et al., 2016).

2.5.1 Intervention

As well as CBNs capturing judgment patterns, Pearl's "Do calculus" has proven to be an effective way of capturing the distinctions in the inferences people make from observations and interventions (Sloman & Lagnado, 2005; Waldmann & Hagmayer, 2005), and more-generally, how they reason counterfactually (Lagnado, Gerstenberg, & Zultan, 2013; Rips, 2010; Rips & Edwards, 2013). The broad idea is that observations license backward inferences while interventions do not. Figure 2.5c gives an example of an inference based on an observation that $X_2 = 1$ (e.g. someone is in a rock band) and Figure 2.5d for inference from an intervention (e.g. making someone join a rock band)

which we write as $\text{Do}[X_1 = 1]$. The greyed arrow and scissors symbol between X_1 and X_2 indicates that the intervention overrides the normal causal relationship. In both cases the probability of X_2's direct and indirect descendants are affected as can be seen by their higher probabilities (green shading) relative to their marginal probabilities shown in Figure 2.2a and b. However, only the observation affects the probability of X_2's normal cause X_1 (e.g. liking rock music), or a knock-on effect on X_1's other effect(s) (e.g. X_9 owning many rock records). The idea that people can imagine virtual interventions helps explain important aspects of thinking. For example, counterfactual "What if..." inferences are often consistent with the idea that people imagine an intervention that makes the counterfactual true with minimal revision of its causal history (Gerstenberg, Bechlivanidis, & Lagnado, 2013; Lagnado et al., 2013; Rips, 2010). Perhaps the most important property of interventions though, is that they allow learners to infer causal rather than correlational structure.

2.6 Active learning

Intervening is a form of *active learning*. Active learning is the study of situations in which learners exert control over the evidence they see (Gureckis & Markant, 2012; Settles, 2012). Controlling incoming information flow is clearly crucial to cognition. There is a massive amount of information available in the world (and in the head) and attention is a limited resource (Lavie, 2005), meaning the cognitive system must be adept at focusing on what is liable to be useful and ignoring the rest. At a low level we constantly exert active control over informational inputs by moving our eyes and our bodies to focus on sources of information relevant to what we are wondering. We see this in striking demonstrations of *change blindness*, in which focusing on one aspect of a scene leads us to miss other major events (Simons & Levin, 1997, and see `https://www.youtube.com/watch?v=IGQmdoK_ZfY` for an example). However, active learning is also key to higher level cognition, in which asking questions, exploring and testing the world play large roles in shaping what evidence human learners receive.

Accordingly, developmental research has shown that even young children are adept at asking useful questions (Lucas, Bridgers, Griffiths, & Gopnik, 2014; Nelson, Divjak, Gudmundsdottir, Martignon, & Meder, 2014; Ruggeri & Lombrozo, 2014). Research with adults has shown that people learn rules (Oaksford & Chater, 1994), categories (Markant & Gureckis, 2010) and spatial concepts (Gureckis & Markant, 2009; Markant & Gureckis, 2012) quicker, and make more accurate classifications (Nelson, McKenzie,

Cottrell, & Sejnowski, 2010) when actively selecting their own samples, as well as being adept at generating informative natural language questions (Rothe, Lake, & Gureckis, 2016). Accordingly, a number of studies have also looked at how people balance exploring an uncertain environment against exploiting what they already know about it. This problem is often tested with *multi-armed bandit problems*, in which participants repeatedly sample from a set of options with unknown, differing and stochastic payoffs, with the goal of maximising long term reward (Macready & Wolpert, 1998; Schulz, Konstantinidis, & Speekenbrink, 2016; Shanks, Tunney, & McCarthy, 2002; Steyvers, Lee, & Wagenmakers, 2009). In this setting, research has found that people explore *more* than they normatively should (Christian & Griffiths, 2016; Tversky & Edwards, 1966), although recent accounts have suggested this could be a rational response to uncertainty about whether payoffs change over time (Speekenbrink & Konstantinidis, 2014). While selecting the right question to ask, place to look, or sample to draw is clearly an essential part of active learning in the world, active learning is particularly important for identifying causal structure.

Because interventions affect things that are causally downstream but not those that are causally upstream, perturbing the world allows causal learners to get data that is otherwise unavailable – information about *causal* rather than merely *correlational* relationships. As an example, suppose you are a medical researcher interested in learning if there is a relationship between X_1 — the presence of some bacterium in the stomach — and X_2 — developing a stomach ulcer. Identifying that patients typically exhibit both or neither of X_1 and X_2 tells you that the two are likely to be causally related but does not tell you in what way. Perhaps the bacterium causes stomach ulcers; perhaps stomach ulcers provide a breeding ground for the bacteria to grow; or perhaps the two phenomena share some other common cause. In the absence of a time cue or pre-existing mechanism knowledge, the direction of causal connections can only be established by performing active *interventions* (experimental manipulations) of the variables. In this example, one might manipulate X_1 by ingesting the bacteria and waiting to see if one develops an ulcer[9], or manipulating X_2 by making cuts in a sample of stomach lining tissue to see if this results in growth of the bacterium. If manipulating X_1 changes X_2 then this is evidence that X_1 is a cause of X_2.

The importance of interventional active learning is well established in education, and

[9]This is, in fact, how Barry Marshall won the 2005 Nobel Prize in physiology or medicine for the discovery that *H pylori* causes stomach ulcers (see Marshall, Armstrong, McGechie, & Glancy, 1985).

developmental psychology, where self-directed "play" is seen as vital to healthy development (e.g. Bruner, Jolly, & Sylva, 1976; Piaget & Valsiner, 1930). Accordingly, a number of developmental psychologists have adopted a "child as scientist" analogy, which sees children as fundamentally engaged in causal hypothesis testing (Gopnik et al., 2004; Gopnik & Sobel, 2000; Sobel & Kushnir, 2006). In adults, a number of studies have found that people benefit from the ability to perform (or watch others performing) interventions during causal learning (Lagnado & Sloman, 2002, 2004, 2006; Schulz, 2001; Sobel & Kushnir, 2006). However fewer have looked at *how* people select what interventions to perform, and these only focus on the case of a single intervention on a single variable in semi-deterministic context (Coenen et al., 2015; Steyvers et al., 2003). In contrast, real world causal learning is generally less constrained, with many variables and probabilistic relationships, requiring extended interactions to reveal the exact patterns of relationships. As I discuss later in the Chapter, spatiotemporal information often provides additional cues to causality, and pre-existing knowledge (or "priors") and mechanism knowledge constrain the space of plausible relationships. However, in the worst case, these other cues are uninformative or unavailable. Many systems propagate too fast to permit observation of time ordering of component activations (e.g. electrical systems); or have hidden mechanisms (e.g. biological systems, psychological processes); or noisy/delayed presentation of variable values (Lagnado & Sloman, 2006); while in others (e.g. crime scene investigation) observations come after the relevant causal process has finished. Furthermore, expectations about causal delays and mechanisms for different causal domains must themselves be learned.

I illustrate this "pure contingency" interventional data in Figure 2.5a. Here a learner intervenes in the 10 variable "world" introduced in Figure 2.2. The learner fixes component X_1 "on" and leaves everything else free to vary.[10] As a result, they observe four of the other nine variables turn "on". Intuitively, this provides some information about the underlying relationships. Assuming we know that this is a generative setting (causes raise the probability of their effects) this intuitively makes X_2 more likely to be a cause of X_4,X_7, X_8 and X_{10}, but less likely to be a cause of X_1, X_3, X_5, X_6, or X_9. However, there is clearly a lot of residual uncertainty. For example it may be that some of the observed activations are indirect or coincidental. Indeed, looking back to Figure 2.2b, we see that only X_4 was directly caused by X_2 in this example. X_7, X_8 were indirect effects

[10]Throughout the thesis I will refer informally to causal systems as being made up of *components*. Where the causal system is described by a CBN, read "components" as equivalent to "variables". Additionally, since I generally focus on binary variables I use "on", *active* or *present* as shorthand for "take the value 1", and "off', "inactive", or "absent" as shorthand for "take the value 0", as appropriate.

and X_{10} must have occurred by chance (in the language of noisy-OR causal models, we can think of it as being caused by factors outside the model).

Fortunately, one can get less ambiguous information about particular edges by controlling variables. The importance of controlled testing is emphasised in education, where "scientific thinking" is taught as a methodology whereby one holds all but the variables related to the relationship of interest constant (Chen & Klahr, 1999; Klahr & Nigam, 2004; Kuhn & Dean, 2005). This is also the basis for *randomised controlled trials* (Cartwright, 1989), where random allocation to conditions averages out the confounding effects of causal influences related to allocation procedure. In the binary-variable, generative-relationship setting, we can "control" by fixing some variables "off" so that they cannot exert confounding causal influences. Figure 2.5b gives an example of a highly controlled test where all but one of the variables are fixed. Intuitively, the fact that X_7 does not turn "on" now is evidence against a direct relationship from X_2 to X_7, which was not apparent from the intervention in Figure 2.5a. Unfortunately, in reality it is not generally possible to control for every conceivable confound at once, nor would it be feasible to test every pair of variables in isolation. Indeed, in this simple case it would take 1024 interventions just to check every direct relationship once. Furthermore, it is clear that by fixing lots of variables, one misses a chance to gather information about relationships between those variables. Thus, the intuition is that the most informative interventions often lie somewhere between maximal open-endedness and maximal control, with the perfect choice dependent on prior expectations and knowledge about the potential structures. This implies that it can take a careful selection of multiple interventions to narrow in on the right causal structure. Fortunately, given a well-defined goal, hypothesis space and prior there is a mathematical answer to which interventions are most useful.

2.6.1 Optimal intervening

In general, we want our interventions to improve our knowledge about the underlying relationships. Thus, we need a way of evaluating possible future knowledge states so we can try to approach those that are more desirable. Choosing queries or experiments that will, in expectancy, maximize some sensible value of one's posterior beliefs is a cornerstone of Bayesian optimal experimental design and decision theory (Good, 1950; Lindley, 1956). Accordingly, there are various proposals in both the mathematical and psychological literature about what is the most appropriate objective function for driving

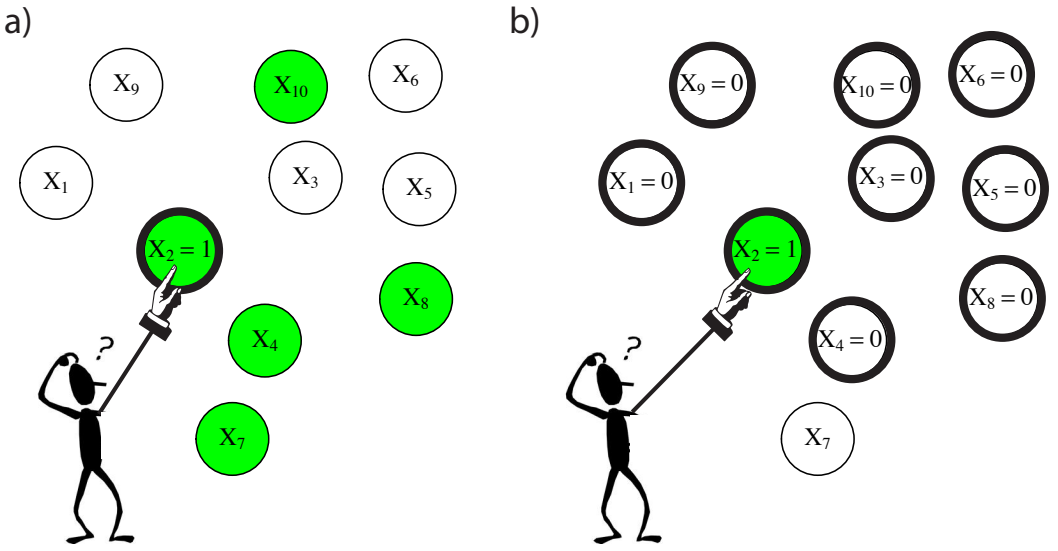


FIGURE 2.5: Interventional learning from covariational evidence; i.e. cases where relationships are probabilistic and outcomes of interventions are all revealed together, green denotes components that turned "on", white denotes those that stayed "off', thick black borders denote variables that are fixed by the learner a) A "simple" intervention b) A "controlled" intervention (extra arms from the learner to the fixed "off" variables omitted for clarity).

active learning in different domains (Butko & Movellan, 2008; Markant & Gureckis, 2012; Meder, Gerstenberg, Hagmayer, & Waldmann, 2010; Nelson, 2005; Nielsen & Nock, 2011; Renyi, 1961; Shannon, 1951; Tsallis, 1988) and I will consider several possibilities in Chapter 3. However, the different objectives generally embody a similar intuition: posteriors that imply greater *certainty* about the true structure are more desirable.

Having settled on an objective, one can go about evaluating different potential interventions through a *preposterior* (Raiffa, 1974) analysis of their possible outcomes and resultant values. To do this, we imagine all the things that could happen if we perform an intervention. For the intervention in Figure 2.5a there are $2^9 = 512$ possible outcomes (all combinations of the 9 free-to-vary variables occurring or not), while for the intervention in Figure 2.5b there are $2^1 = 2$ (X_3 will either turn on or not). In principle, we can calculate what our posterior beliefs would be after each of these potential outcomes. Using our objective function, we could then calculate the value of each of these possible future posteriors. By weighting each value by the marginal likelihood of observing that outcome and averaging, we end up with an expected value for the intervention. By doing this for all possible interventions (i.e. each possible setting of variables "on", "off" or "free to vary") we can choose whichever setting has the maximum expected value. As if this were not intensive enough, strictly we should also plan ahead, assuming we will also choose all future interventions optimally, and choose the intervention that gets us

to the greatest rewards on average by the end of learning (Puterman, 2009). Returning to Figure 2.1, I depict this process as a loop, where intervention choices are guided by a combination of preposterior analysis given current state of knowledge about the true causal model (prior), and the "utilities" that determine how valuable potential future knowledge states are to the learner. Each chosen intervention affects the data produced by the world (e.g. the learner now gets to see the actual outcome) allowing them to improve their model and repeat the process. I look in detail at the utilities guiding intervention selection in Chapter 3.

The few papers that have analysed intervention choice have generally shown that human interveners select tests that are informative, generally more so than behaving randomly, yet display systematic suboptimalities and biases. I describe this behavioural evidence in more detail in motivating my account of bounded intervention selection below.

2.7 Intractability

I have now introduced CBNs as a useful framework for modelling people's causal structure representations, Bayesian inference as a framework for modelling optimal inference, and preposterior analysis for modelling optimal active learning. However, the number of possible CBNs grows very rapidly with the number of variables (Table 2.1). So too does the number of potential interventions and the number of outcomes to be averaged across for each intervention. In principle, all combinations of potential model, intervention and outcome should be considered in order to select the most valuable intervention in expectation. This means that inference and choosing interventions scale so poorly in the number of variables, they are fundamentally intractable for any plausibly bounded learner (Cooper, 1990; van Rooij et al., 2014).

2.7.1 Heuristics

One response to the divergence between the impossible demands of optimal calculation and limited cognition is to step away from the desiderata of relating cognitive processes to rationality. Herbert Simon (1956) famously argued that, in situations where the optimal solution is intractable, one should satisfice rather than optimise. A portmanteau of "satisfy" and "suffice", satisficing means searching through options until finding one that exceeds some prechosen acceptability threshold. Building on this idea, Simon formalized the idea of bounded rationality, as behaviour that is rational given one's limitations

Table 2.1: The Number of Possible Structures, Interventions and Outcomes For 1–10 Binary Variables

Variables	Structures	Interventions	Outcomes
1	1	3	1
2	3	9	2
3	25	27	4
4	543	81	8
5	29281	243	16
6	3781503	729	32
7	1138779265	2187	64
8	783702329343	6561	128
9	∼1213442000000000	19683	256
10	∼4175099000000000000	59049	512

Note: "Structures" gives the number of directed acyclic graphs. "Interventions" gives the number of combinations of fixed "on" fixed "off" and free variables. "Outcomes" gives the maximum number of outcomes i.e. all combinations of "on" and "off" assuming a single variable is intervened upon.

(e.g. in processing power, storage and or time), arguing that it is very hard to determine optimal bounded behaviour. Indeed, calculating the optimal amount of cognitive resources to devote to a problem, can be a lot harder than devoting those resources to the problem itself (Chow, Robbins, & Siegmund, 1971). Simon argues that this means cognition simply finds or evolves, approaches that are "good enough" for purpose, and sticks with them unless their performance declines or a better strategy is discovered.

Building on Simon's ideas, cognitive and computational modelling has identified a wealth of heuristics, or "computational shortcuts", that describe how people behave, often surprisingly well, in circumstances where shooting for optimality seems unachievable (Gigerenzer, 2001; Gigerenzer, Todd, & ABC Research Group, 1999; Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1975). In the context of causal learning, proposed heuristics have included treating causal relationships as independent so as to learn at the level of individual relationships rather than overall models (Fernbach & Sloman, 2009; Waldmann, Cheng, Hagmayer, & Blaisdell, 2008), and making simplifying assumptions about the functional form of model, e.g. treating them as deterministic or near-deterministic (Lu et al., 2008; Mayrhofer & Waldmann, 2016).

By design, proposed heuristics tend to do well in particular environments, and provide good descriptions of human behaviour in specific tasks. Unfortunately, there are no free lunches when it comes to computation (Wolpert & Macready, 1997). The effectiveness of a computational shortcut always depends on its fit to the environment. Heuristics can be highly effective when they pick up on useful stable properties of an environment that make more complex computations unnecessary, but can be worse than chance when

applied in inverted or antagonistic environments. As cases in point, we will see in Chapter 3 that a simple heuristic ignoring dependencies between causal relationships can be near-optimal for problems where the causal relationships are in fact quite sparse (most variables are unrelated). However, as Figure 2.5a suggests, doing this naïvely in situations with many relationships can lead to unacceptably many false positives, such as inferring spurious direct connections from X_2 to X_7 and X_8. Likewise, simplifying assumptions about functional form are effective when they are not far off the mark — e.g. if the true relationships (or those worth storing) are actually near enough deterministic — but misleading if the causal relationships are actually functionally very different.

In general, people often seem to behave consistently with heuristics when conditions favour them, yet, when sufficiently pushed, often reveal that they are capable of greater sophistication (e.g. Newell & Shanks, 2003, 2004; Shanks & Lagnado, 2000). Thus, a looming meta-problem for a heuristic theory of cognition is how people choose which heuristics to apply when, a problem that threatens to reinstate the complexity that heuristics are supposed to avoid (Cooper, 2000; Newell & Shanks, 2007; Parpart, Jones, & Love, in revision).

2.7.2 Rational process models

Despite lacking a formal relationship with optimality, there is a powerful idea at the heart of the heuristics research program. There are often much cheaper ways of interacting effectively with natural environments than the use of maximally complex probabilistic models, and bounded agents must strike a balance between internal computation costs and the costs of suboptimal behaviour. A recent movement is to treat heuristics as components of potentially rational approximations (Griffiths et al., 2015; Lieder & Griffiths, 2015; Lieder, Griffiths, Huys, & Goodman, under review; Parpart et al., in revision). This approach uses the mathematics and algorithms of principled approximation to explore frugal strategies while keeping track of their formal relationship with optimality (Sanborn, Griffiths, & Navarro, 2010).

Machine learning is a field that has had to take approximating probabilistic inference very seriously (Bishop, 2006). Accordingly, it has seen the development of a range of approaches for approximating intractable computations. In general these approaches

provide ways of trading off accuracy against computational efficiency and storage requirements, providing a space of algorithms within which the ideal trade-off exists.[11] Accordingly, a recent approach is to treat the approximation methods developed in machine learning as (in strong form) *candidate* process level models, or (weaker) *inspiration* for process level models. In this thesis I take inspiration from rational sampling approximations in developing my model of bounded causal inference, and explore a range of approximations and heuristics for modelling intervention selection.

Sampling approximations to inference

Sampling approaches are a common family of approximate algorithms. These use stochastically generated samples to represent intractable distributions, with methods for generating and weighting these samples asymptotically approximating Bayesian inferences. Sampling methods have levels of accuracy that typically depend on the number of samples, while different sampling schemes are more efficient for different types of problem.

There have been a number of proposals that the brain uses sampling to approximate probabilistic inference (e.g. Hamrick, Smith, Griffiths, & Vul, 2015; Sanborn, 2015; Stewart, Chater, & Brown, 2006). A strong version of this proposal is that the brain never calculates probabilities explicitly, but conforms approximately to probability theory in virtue of being a powerful general purpose Bayesian sampler (Sanborn & Chater, 2016). To make sense of the relevance of sampling approximations for causal cognition, and the types of behavioural phenomena they predict, I briefly review three common sampling schemes.

Simple Monte Carlo sampling A common problem is comparing models without exact knowledge of their parameters. To be Bayesian, we should treat uncertain model parameters as random variables. As a minimal case, suppose we are simply interested in determining whether X_1 causes X_2, so our hypothesis space is simply $\mathcal{M} = \{m_1 = [X_1 \nleftrightarrow X_2], m_2 = [X_1 \rightarrow X_2]\}$ (see Figure 2.7a). Suppose also that we have intervened on X_1 three times, switching it on twice and off once. Each time X_2 did the same thing as X_1. This is our data $\mathbf{d}$. Intuitively, this favours m_2 over m_1, but to what extent? How sure should we be, based on these three tests? In the figure, I assume

[11] Although, of course some approximation strategies are more efficient than others, meaning that a different approach might always offer a better trade-off.

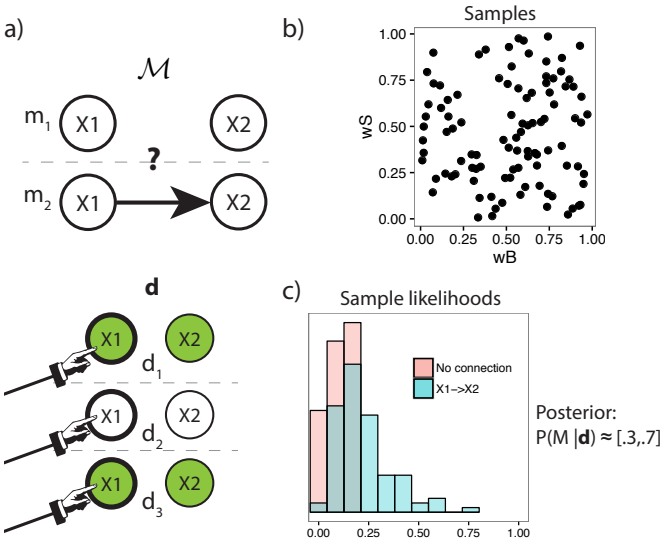


FIGURE 2.6: Using Simple Monte Carlo sampling to estimate the probability of a connection between X_1 and X_2. a) Hypothesis space $\mathcal{M}$ and data $\mathbf{d}$. b) One-hundred uniformly distributed random causal *strength* and spontaneous activation *base rate* samples. c) The sample likelihoods and resulting posterior in which $m_2 : [X_1 \rightarrow X_2]$ is now more likely than $m_1 : [X_1 \nrightarrow X_2]$.

the true model has a noisy-OR parametrisation, but we do not know the exact *strength* w_S of the putative connection or the *base-rate* w_B with which X_2 turns "on" by itself. I also assume we start from complete ignorance about the true model, strength and base rate (our priors $P(M)$ and $P(w_S, w_B)$ are uniform). To approximate the model posterior $P(M|\mathbf{d}) = \int P(M|\mathbf{d}, w_S, w_B)\ \mathrm{d}w_S\ \mathrm{d}w_B$, we can draw (paired) samples from our parameter prior $P(w_B, w_S)$ (Figure 2.7b), calculate the likelihood of the data with each, then average. For samples where the expected relationship is strong and base rate is low (e.g. near the top left of Figure 2.6b) these data are much more likely under m_2 because the connection is needed to explain X_2's activations. For samples of low strength and high base rate (e.g. near the bottom right of Figure 2.6b) the data is similarly likely under both models, since both predict X_2 will occur frequently by chance with little influence from X_1. Thus, by averaging we are performing an approximate "numerical integration" over our parameter uncertainty, finding that we should update our beliefs a moderate amount toward m_2 as shown in Figure 2.7c.

This strategy is natural for simple cases, however more sophisticated schemes are needed for maintaining probabilistic beliefs as evidence arrives sequentially, and for drawing samples from intractable distributions, such as a distribution over possible causal models.

Particle filtering One common approximation for situations where evidence arrives sequentially, is to maintain a manageable number of individual sample hypotheses, or

"particles" (Liu & Chen, 1998), with weights corresponding to their relative likelihoods. The ensemble of particles then acts as an approximation to the desired distribution. Sophisticated reweighting and resampling schemes can then filter the ensemble as data are observed, approximating Bayesian inference. In our example, a histogram over our w_B and w_S samples approximates the probability distribution over their possible values. This histogram would be approximately flat in the case of the uniform prior we assumed above, but must change shape if it is to represent accumulating knowledge about the parameters. In Figure 2.7a, I extend 2.6 to give a basic example of this idea. To keep track of our evolving beliefs about w_S and w_B as we observe more data, we can resample with replacement from our set of "particles" with probabilities determined by the likelihood they assign to the latest datum. Thus the particles pile up in regions of high probability, with the histogram approximating the Bayesian posterior. In the Figure, we see the filtering reflects a rational increase in preference for high strength parameter and low base rate parameter as a learner tests a (known) $X_1 \rightarrow X_2$ relationship five times and finds it quite reliable. The filtering procedure reduces the diversity of the samples over time because each resampling step will tend to clone some existing particles while letting others go extinct. Thus, there are various techniques for "rejuvenating" the set of particles — i.e. generating new ones without affecting the distribution (Chopin, 2002).

Reliance on filtering a limited number of samples has been proposed as an explanation for a number of human biases in behavioural tasks. For example, under certain types of filtering, samples consistent with early observations can dominate and lead to a failure to generate any samples in regions more consistent with later samples, leading to primacy (Abbott & Griffiths, 2011; Levy, Reali, & Griffiths, 2009; Sanborn et al., 2010).

Additionally, particle filtering has been used as an explanation for individual variability. A pervasive phenomenon in human decision making tasks is probability matching (Myers, 2014; Shanks et al., 2002; Vulkan, 2000) where, rather than always choosing the best option, responses appear to be chosen in proportion to their probability of being the best. This behaviour seems strange at first glance, but is more intelligible if we suppose that people make decisions based on very limited number of posterior samples (Brown & Steyvers, 2009). In fact, in associative learning (Courville & Daw, 2007), categorization (Sanborn et al., 2010) and binary decision making (Vul, Goodman, Griffiths, & Tenenbaum, 2009), it has been proposed that people's beliefs actually behave most like a single particle. Thus one plausible approximation for structure learning is to consider just a few or even a single structural hypotheses at a time, in place of the full distribution. One proposal for the problem of causal structure learning, which I review extensively

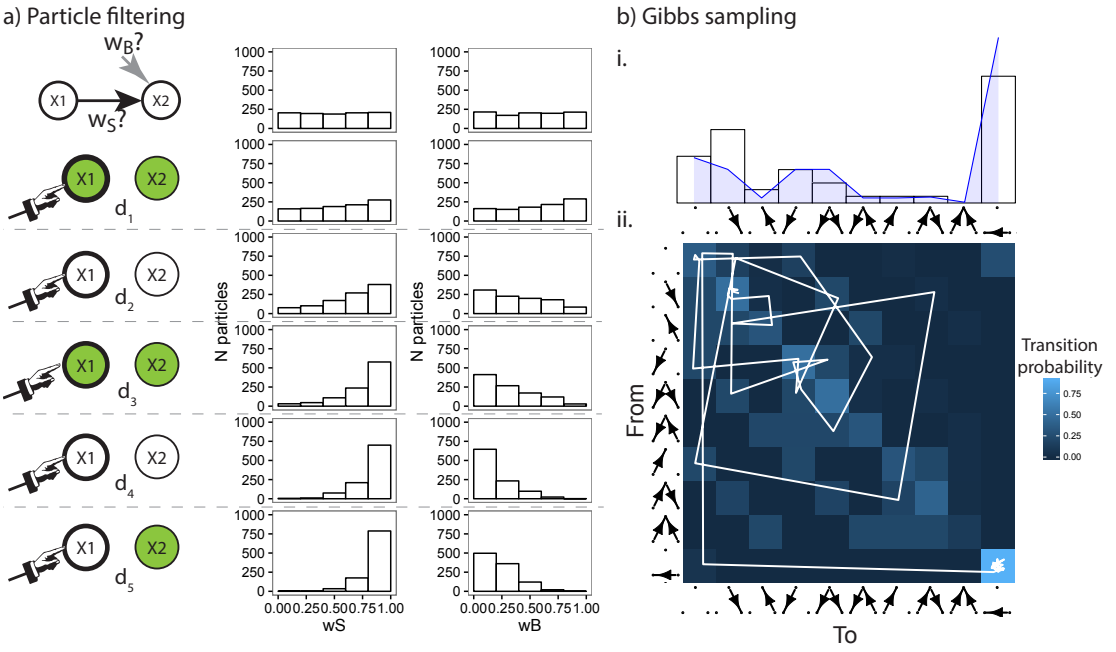


FIGURE 2.7: Two examples of sample based approximate causal inference. a) Particle filtering 1000 paired $w_S^{(i)}$ and $w_B^{(i)}$ samples, approximating an evolving posterior on $p(w_S, w_B)$. After each datum, particles are resampled (with replacement) with sample weights proportional to $P(\mathbf{d}|w_S^{(i)}, w_B^{(i)}, m)$. b) Gibbs sampling in causal model space. i. Histogram shows locations visited, blue overlay shows true posterior probability of these models. ii. Matrix shows probability of transitioning from the model corresponding to the row, to the model corresponding to the column. All transitions resample one edge at a time conditional on the current state of the others. White line shows the sampled path that generated the histogram.

in Chapter 5, is *win-stay, lose-sample* (Bonawitz, Denison, Gopnik, & Griffiths, 2014), which is the idea that learners generate a single sample hypothesis which they keep as their candidate until they observe strongly refuting evidence whereupon they resample a new candidate from the posterior.

Reliance on a limited number of candidate model "particles", or even a single candidate, seems like an important idea for understanding bounded structure inference. However, as the number of particles reduces, the degree of approximation increases. For a single particle, the filtering approach is degenerate (the distribution is represented by a 1 bar histogram). Thus, it becomes very important to have a method for resampling hypotheses from the posterior.

Markov Chain Monte Carlo Another class of useful machine learning methods, Markov Chain Monte Carlo (MCMC) sampling, involves generating sequences of hypotheses, each linked to the next via a stochastic transition mechanism which asymptotically approximates the posterior distribution. Under various conditions, this implies

that the sequences of autocorrelated sample hypotheses form a Markov chain with a stationary distribution that is the full, intended, posterior distribution (Metropolis, Rosenbluth, Rosenbluth, Teller, & Teller, 1953). The samples will appear to "walk" randomly around space of possibilities, tending to visit more probable hypotheses more frequently. If samples are extracted from the sequence after a sufficiently long initial, so-called burn-in, period, and sufficiently far apart (to reduce the effect of dependence), they can provide a good approximation to the true posterior distribution.

MCMC sampling has been proposed as an explanation for several behavioural phenomena. For example Lieder, Griffiths, and Goodman (2012) propose that MCMC sampling can explain anchoring effects, in which judgments are dependent on an in-principle-irrelevant initially provided value. This is based on the assumption that processing constraints limit the length of the sampling chains so that they retain dependence on their starting point. Limited MCMC chains have also been proposed as an explanation for unpacking effects (Dasgupta, Schulz, & Gershman, 2016), where conjunctive categories are judged as more or less likely than the union of their "unpacked" members, as well as predicting transition times between percepts (Gershman, Vul, & Tenenbaum, 2012). Thus, another potential aspect of rational approximation in causal learning could be stochastic search in causal model space.

There are typically many different classes of Markov chain transitions that share the same stationary distribution, but differ in the properties of burn-in and subsampling. A simple form of MCMC, that is naturally applicable to structure inference is Gibbs sampling (Geman & Geman, 1984). Informally, Gibbs sampling works by resampling one part of a large multivariate distribution at a time. In the current context, these might mean resampling individual edges in a causal model. Figure 2.7b gives an example of this. The blue shaded region in Figure 2.7b i. shows a desired posterior distribution over 10 models. Figure 2.7b ii. shows the transition probabilities for the sampling chain. The shading in each cell shows the probability of an MCMC move from the model in the row to the model in the column, and the white line gives an example search path of length 50. For example, the top row of the matrix shows the probability of transitioning from the unconnected model to any of the other models. The only models that are accessible are those that can be reached by changing a single edge. Returning to the histogram above, we see that the number of times the path visits each model is indeed approximately proportional to its posterior probability. I will develop a process model of the generation of candidate hypotheses based on such a "local" Gibbs-style search in Chapter 5.

2.7.3 Bounded intervention selection

In situations where a posterior is already hard to evaluate, calculating the globally most informative intervention will almost always be infeasible. Fortunately, a variety of methods have been developed that allow tests to be selected that are more useful than random selection, but do not require the full expected information gain be computed (Settles, 2012). In general, these approaches are heuristic in that they generally have limited guarantees about the circumstances in which they will be effective.

Many of the active machine learning heuristics rely on the current, rather than expected uncertainty (e.g. *uncertainty sampling* which chooses based on outcome uncertainty under the prior). However others use the predictions under just a few favoured hypotheses (e.g. *query by committee*) as a substitute for the full expectancy calculation. The former strategy relies on maintaining a complete prior distribution, which I have already suggested is implausible in the case of causal structure inference. Therefore in this thesis, I focus on developing a general purpose approach based on the latter idea, that people consider only a small set of hypotheses when choosing interventions.

We see this idea in Markant, Settles, and Gureckis (2015), who propose that self directed learning is better understood as favouring local rather than global uncertainty, e.g. uncertainty pertaining to a single dimension of the problem at a time. However, this raises the metaproblem of which of the large set of possible local partitions of the hypothesis space people focus on when choosing tests. In this thesis I develop the proposal by considering a variety of forms of local focus, in the process re-framing commonly proposed heuristics for query selection. In particular, I propose that *constraint seeking* (Nelson et al., 2014; Ruggeri & Lombrozo, 2014), and *confirmatory testing* (Klayman & Ha, 1989; Nickerson, 1998) are complementary types of "local focus" which learners switch between.

Constraint seeking or local testing?

A commonly proposed heuristic for efficient search in deterministic domains is to ask about the dimension that best divides the hypothesis space, eliminating the greatest possible number of options on average. This has been called "constraint-seeking" (Ruggeri & Lombrozo, 2014) and the "split-half heuristic" (Nelson et al., 2014). For example, in the children's game "Guess who" one player chooses one of a set of cards with different faces on the back. Their opponent then asks binary questions with the goal of

identifying the chosen card. In this setting intuitively good questions are those that split the remaining options in half. For instance if the cards are gender balanced, asking "is the person female?", will eliminate about half of the options on average. Bad questions will pick out features that are very unbalanced, eliminating fewer cards on average. Constraint seeking corresponds to the most valuable question for a wide range deterministic environments and greedy objectives. The case of using interventions to reveal causal structure is rather more complex but similar principles apply if one makes simplifying assumptions. One source of complexity is that I have assumed causal models are not necessarily deterministic, meaning that the outcome of an intervention does not rule structures out but rather renders them more or less likely. However, as already mentioned, people may adopt a presumption of (near) determinism when reasoning in complex causal models (e.g. Lu et al., 2008; Mayrhofer & Waldmann, 2016), suggesting that they might do the same when selecting tests. Another complexity comes from the fact that in general interventions are not binary questions. This is evident in Figure 2.8a, where there are $2^2 = 4$ possible outcomes of turning X_1 "on". In principle, an intervention with 4 possible outcomes could split the hypothesis space into 4 equal parts, leaving you with only 6 or 7 possibilities after your first intervention. However, the split is not that even, with the most likely outcome being Option 1 (nothing happens) leaving you with 12 remaining possibilities. Thus by performing this intervention you can expect to rule out all but ≈ 8 models on average. It turns out that the best split in deterministic settings is achieved by querying the effects of a single randomly chosen variable, essentially asking: which other variables (if any) are descendants of variable X_i in the true model?

For comparison, Figure 2.8b shows a more controlled test that fixes X_3 "off" so focusing on the putative $X_1 \rightarrow X_2$ relationship. While this rules out fewer hypotheses on average (leaving ≈ 14 in expectation), it is intuitively much easier to interpret as there are only two outcomes and both provide unambiguous "local" information about this relationship (indeed it is necessary to perform such an intervention to distinguish certain models). I will consider this trade-off in detail in Chapter 5, and revisit it throughout the thesis.

Confirmatory testing

Another commonly proposed strategy is confirmatory evidence gathering, in which learners choose tests that attempt to confirm or dis-confirm their current hypothesis rather

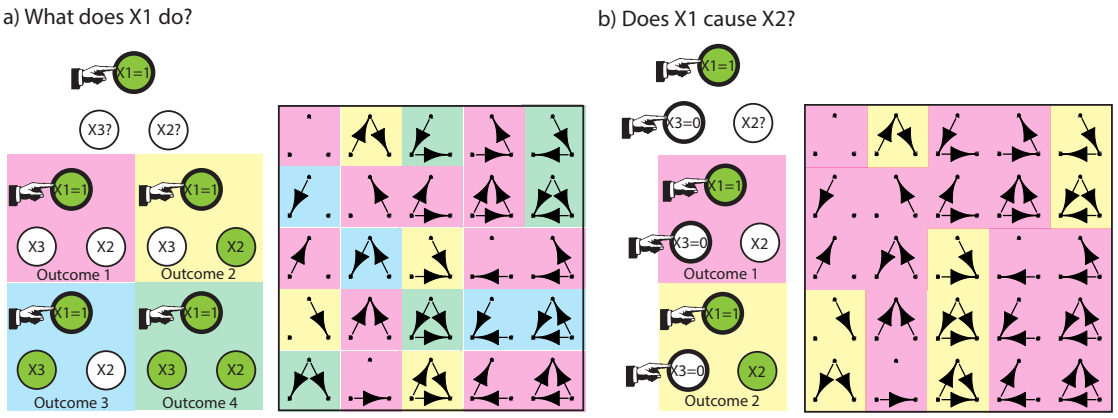


FIGURE 2.8: What can you learn from interventions in deterministic three variable problems? a) An intervention that just fixes X_1 has four possible outcomes, each eliminating all but the subset of the models corresponding to the colour. b) An intervention that fixes X_1 "on" and X_3 "off" has two possible outcomes, depending whether $X_1 \rightarrow X_2$.

than to reduce uncertainty over the whole set of options. Confirmatory evidence gathering appears to be a ubiquitous psychological phenomenon (Nickerson, 1998). Although confirmation seeking is widely touted as a bias, it can also be shown to be optimal, for example under deterministic or sparse hypotheses spaces or peaked priors (Austerweil & Griffiths, 2011; Klayman & Ha, 1989; Navarro & Perfors, 2011).

In the context of causal structure learning, we see this in Steyvers et al (2003) who propose a related *rational test model* that selects interventions with a goal of distinguishing the current most probable hypothesis from a null hypothesis that there are no causal connections, ignoring the possibility of the true structure being something else entirely.

2.8 Beyond Bayesian networks

I have so far discussed causal learning within CBN framework, treating the problem as directly analogous to learning CBNs by variable-setting. While this gets at the heart of the problems of interventional causal learning, it is also certainly a major simplification. The CBN framework alone is not an adequate model of causal representation, largely because it is a very limited representation of physical causality. CBNs say nothing about how causal relationships play out in time and space. Furthermore, variable setting is a very limited characterisation of intervening in the causal world. The human experience of causality is much richer, with relationships embedded in a spatially and temporally continuous natural world, meaning that we often have rich knowledge about what causal relationships are plausible for familiar entities, how long different causal processes take

and even "mechanistic" knowledge about how causal relationships play out in space. An adequate conception of human causal representation and active learning must have space for these richer representations as well as explaining how they can constrain and enrich causal inference.

2.8.1 Richer functional forms

The obvious way to enrich the CBN framework is to define relationships in terms of more than just probabilistic dependence. For instance, we can imagine imbuing causal relationships with parametric form in terms of their spatiotemporal extension as well as their probabilistic strength. For instance a rich representation of the relationship between playing in a rock band and developing tinnitus might encode both the probability of the relationship obtaining at all in a given case, but also the expected timeline of tinnitus onset in both non-rock-band-members and rock-band-members. In this way, the time of tinnitus onset can be informative about whether the rock-band playing was a likely cause. We might go even further and consider the form of causal relationships in space. For example, a doctor might have a mechanistic understanding of how sound waves cause damage to ears and how this impacts on auditory signals in the brain. Again, we can imagine encoding this in our causal model so that we can reason about potential interventions (e.g. that might reduce symptoms). So how can we represent these rich spatiotemporal details? We are often capable of reasoning at different levels of abstraction. For instance, we can talk about a loose causal relationship between playing in a rock band and developing tinnitus, without worrying about the fine-grained mechanistic details of this relationship. Yet, at the same time we are also capable of reasoning downwards (e.g. explaining the high level causal relationship in terms of fine-grained mechanistic details) and upwards (e.g. abstracting high level causal relationships from fine-grained mechanistic details). Thus, a full account of causal cognition must account for how we are able to perform inferences spanning levels of abstraction and generality.

2.8.2 Intuitive theories, and hierarchical representations

One approach that tries to account for people's situational and mechanistic knowledge within the probabilistic inference framework is the idea that people form *intuitive theories* about how different domains in the world work (Gerstenberg & Tenenbaum, to

appear; Griffiths & Tenenbaum, 2009; Lake, Salakhutdinov, & Tenenbaum, 2015; Tenenbaum, Griffiths, & Niyogi, 2007). These are modelled as hierarchical priors (or probabilistic programs)[12] that can be used to rapidly generate and evaluate hypotheses during future encounters with a domain. The thought is that people organise their causal general knowledge hierarchically, with the core abstract features of causation at the top and increasingly domain- and context-specific features below. Each level of the theory defines a probability distribution at the level below. As a simple example, we have high-level general knowledge about the kinds of causal relationships that obtain in medical domains. For instance we might know that behaviours can cause diseases, and diseases can cause symptoms but that causality cannot run the other way (e.g. from diseases to symptoms or symptoms to behaviours). We can represent this high level knowledge with a *class graph* (see Figure 2.9a) constraining the plausible causal relationships to be those that run in the right direction between these classes of variables. This theory can be used to constrain inference to domain-consistent hypotheses (e.g. Figure 2.9b) . In this way, a learner can make use of their domain knowledge to learn quicker in subsequent encounters with medical domains. In the theory-based causal inference framework, a learner's world knowledge gets richer, their causal judgments can rely more strongly on identification of the current domain and application of the appropriate domain-specific knowledge, meaning they need not start from scratch each time. This framework makes space for domain representations with rich spatiotemporal functional forms that further constrain the space of plausible models. If we know how long different symptoms take to present, how close contact needs to be for a virus to transfer between people and so on, we can use this information to further accelerate inference.

Another idea about the vertical structure of causal representation appears in philosophy of science where *recursive Bayesian networks* have been proposed as models of reductive and mechanistic explanation (Casini, Illari, Russo, & Williamson, 2011; Clarke, Leuridan, & Williamson, 2013; Williamson & Gabbay, 2005). In a recursive Bayesian network, nodes and edges can contain their own, smaller, Bayesian networks made up of smaller parts, which can in turn contain networks of still-smaller parts (see Figure 2.9b for an example). The behaviour of the level below defines the functional form of the level above. In this way we might *explain* the probabilistic causal relationship between rock-band-playing and tinnitus by referring to a more detailed and mechanistic account of the relationship in terms of smaller internal parts like sound waves and ear drums,

[12]Probabilistic programs use stochastic functions to generate distributional outcomes (Goodman, 2013). For example, a probabilistic program might stochastically generate nodes and connections between nodes to generate causal model hypotheses on the fly, where on average this is equivalent to sampling these causal models from a desired distribution over all possibilities.

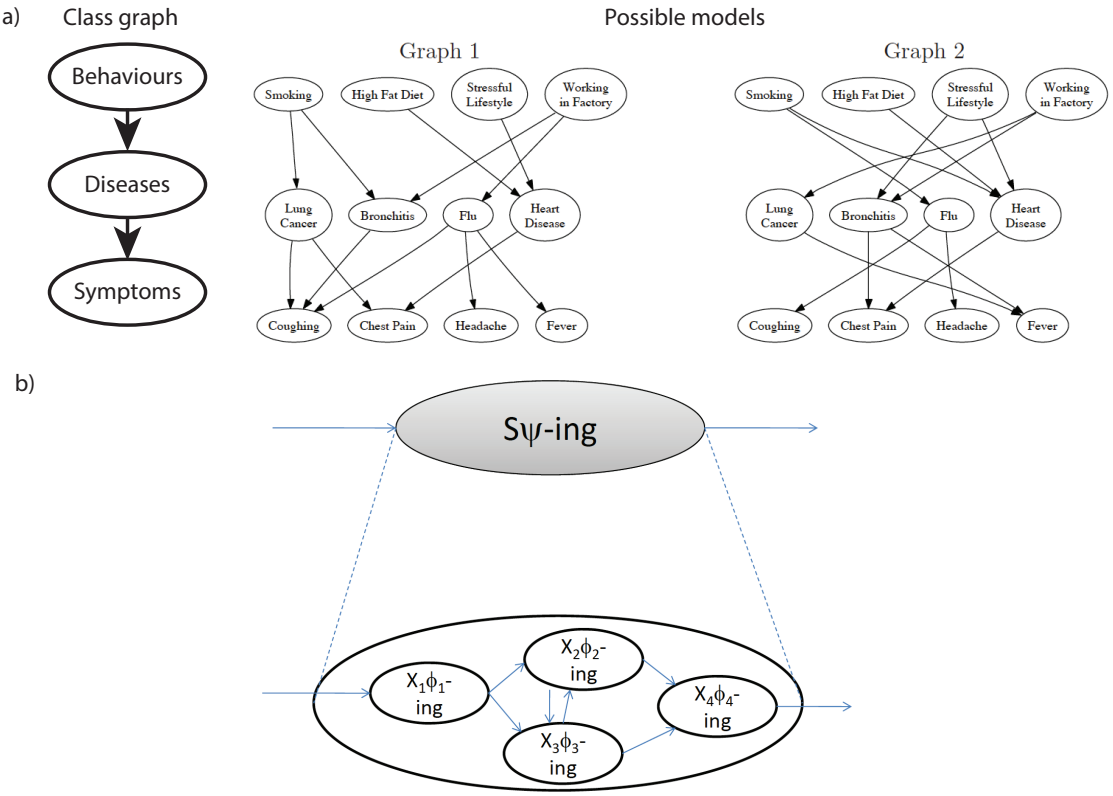


FIGURE 2.9: a) A graph at the level classes of event (Behaviours, Diseases and Symptoms) constrains the hypothesis space of models at the level of particular events. Reproduced with permission from Tenenbaum et al. (2007) b) An element of a recursive Bayesian network. At a high level phenomenon S plays a role in a high level causal representation. Within S there is a low level mechanism made up of parts $X_{1:4}$, that explains the behaviour of S. Reproduced with permission from Clarke et al. (2013).

that overall give rise to the higher level relationship. Equally, we might abstract from the complex dynamics of sound waves, ear drums and neural signals a simple compound concept of tinnitus. Both approaches to formalising layered beliefs highlight the "blessing of abstraction" (Gershman, 2016; Goodman, Ullman, & Tenenbaum, 2011) — which is the idea that abstraction from the particular to the general is an important and early-appearing aspect of the successful development of complex cognition.

While neither hierarchical organisation nor reductive explanation are core topics of this thesis, it is important to keep in mind the ways in which human causal representations, and correspondingly interventional causal learning, is richer and more multilayered than can be fully captured within the CBN (or any single) framework. Thus, in the later chapters of the thesis I focus on the role of elements of richer representations, and correspondingly richer notions of intervention and evidence on causal learning.

2.8.3 The role of time

Prima facie, time is very important to both causal learning and representation. Since, by definition, causes precede their effects (Hume, 1740), the order of events are an obvious cue to causal directionality. Furthermore, if people have expectations about how long causal relationships take to work, or how reliable cause–effect delays will be (e.g. in a familiar domain), this can provide rich additional information. People have been shown to make systematic use of temporal information in inferences about a single putative relationship (Buehner & May, 2003, 2004; Buehner & McGregor, 2006), and causal beliefs also influence time perception (Bechlivanidis & Lagnado, 2013; Buehner & Humphreys, 2009; Haggard, Clark, & Kalogeras, 2002). Indeed, when temporal cues have been pitted against statistical cues experimentally, causal judgments have tended to be dominated by temporal information (Burns & McCormack, 2009; Frosch, McCormack, Lagnado, & Burns, 2012; Lagnado & Sloman, 2004, 2006; Schlottmann, 1999). However the role of time in causal learning over multiple variables has not been investigated systematically, and I do this in Chapter 6 and 7.

CBNs are very limited in their representation of time. The interventional calculus embodies the minimal assumption that causes precede their effects, but it says nothing about the relative delays of competing causal pathways. Worse, by being based on factorisation of a joint probability distribution, CBNs cannot represent relationships that form loops or cycles. In contrast, such dynamic relationships seem common in the world — e.g. that liking rock music might make you join a rock band which might further increase your enjoyment of rock music. All sorts of real world processes, from population change (Malthus, 1888) to economic, biological and physical processes are characterised in terms of reciprocal and dynamical causal processes. In experiments, people frequently report causal beliefs that include cyclic relationships when allowed to do so (Kim & Ahn, 2002; Nikolic & Lagnado, 2015; Sloman, Love, & Ahn, 1998). While there are ways of adapting the CBN formalism to capture cycles — (e.g. Dean & Kanazawa, 1989; Lauritzen & Richardson, 2002) and see Rehder (2016) for a recent review — none of these proposals capture how cause–effect relationships unfold in continuous time, where some relationships might occur much faster or slower than others.[13]

In Chapter 6, I will develop a time-extensive representational framework that supports cyclic relationships and encodes expectations about causal delays. The basic idea is that we can represent causal relationships in terms of delay distributions as well as

[13] Although see Pacer and Griffiths (2011, 2015) for a model that covers related, rate-based, cases.

contingencies, so that representations support predictions about *when* future events will occur, have occurred, or would occur following interventions, as well as allowing observed event timings to provide additional data about the true causal model.

Time is also intimately related to the idea of intervention. Lagnado and Sloman (2004) propose that interventions (in real time) act like a strong order cue, with events happening shortly thereafter liable to be associated causally with the action. This also has a relationship with the idea of operant conditioning (Skinner, 1963). However the focus on contingency data in the causal learning literature, and on non-causal learning in the conditioning literature has led to the question of how people select interventions and interpret their outcomes in continuous time to remain relatively unexplored. I look at this in Chapter 7.

2.8.4 The role of space

Stepping further beyond the covariation-based accounts of causal inference embodied by the CBN approach, a number of psychologists and philosophers have emphasised the role of mechanism knowledge in causal judgments (Ahn, Kalish, Medin, & Gelman, 1995; Craver, 2006; Illari & Williamson, 2012; Salmon, 1994; Shultz, 1982; Waskan, 2006; White, 1995; Williamson, 2004). Here, the idea is that people's causal judgments are often grounded in understanding of the mechanisms and processes involved. For example, Ahn et al. (1995) find that participants tend to provide mechanistic causal explanations and are more convinced by evidence that includes mechanism information.

One area in which causal judgments often seem strongly based in mechanistic knowledge is the domain of physical causality. Psychologists have long been interested in how people make judgments about the physical world, such as how objects will move when thrown or dropped (e.g. McCloskey, Caramazza, & Green, 1980), what will happen when they collide and exert forces on one another (Michotte, 1946/1963; White, 2009; Wolff & Barbey, 2015; Wolff & Shepard, 2013). In general, early work on intuitive physics emphasised people's failures, finding systematic deviations between the dictates of Newton's laws of motion and people's judgments about the trajectories of objects in simple scenes. For instance, many people predict that an object dropped from a moving source will fall straight downward — it will actually continue forward and accelerate downward in a parabola. However, more recent work in the probabilistic inference tradition, has suggested that human successes in predicting physical scenes are more impressive than

their failures. A key finding is that even young children seem to have strong expectations about object permanence and contiguity (Baillargeon, Spelke, & Wasserman, 1985). Battaglia, Hamrick, and Tenenbaum (2013) show that people can make accurate judgments about stability of stacks of objects and Smith and Vul (2014) show they can make accurate forward and backward predictions about object motion. They have success modelling judgments under the assumption that the participants have rich enough intuitive theories of physics to mentally simulate simple physical scenes, much like a computer game graphics engine (e.g. Tenenbaum et al., 2011). From this perspective, some of the biases observed by McCloskey et al might result from rational uncertainty about the physical quantities involved — i.e. masses of the objects, friction, air resistance etc (Hamrick, Battaglia, Griffiths, & Tenenbaum, 2016; Hamrick et al., 2015; Sanborn, Mansinghka, & Griffiths, 2013). In line with this, Gerstenberg, Goodman, Lagnado, and Tenenbaum (2015) show that people's causal and counterfactual judgments reflect the robustness of physical causal processes. For instance, an expert striker who scores convincingly in a football game is judged as more responsible for the outcome than an amateur who trips over and accidentally kicks the ball in the goal (Lagnado et al., 2013). These findings are consistent with the idea that people make judgments of causal responsibility using internal simulations, and are more uncertain when these simulations are not robust to perturbations (e.g. in starting conditions).

When learning about causal relationships in the physical world, observations combined with an intuitive theory of physical causality can already provide rich scope for inference. For example, if we observe two objects colliding, the paths of the two objects in combination with Newton's laws of motion, provide lots of information about whether one or other object is heavier, or more elastic. Thus, in principle, one can work back from known physical laws to infer the latent, causally relevant properties of physical relata. However, it is clear that these properties need to be revealed through dynamics, that the right, revealing, dynamics might rarely occur by chance, or close or clear enough to be perceived reliably. As a simple example, it is hard to tell how heavy something will be until you try picking it up, although observing how it fares when other objects interact with it can help. Thus, a richer notion of intervention captures the idea that we often prod and poke at the physical world in ways that we hope make these latent properties reveal themselves. This means creating situations that provide a lot of information about properties of interest while minimising confounding information. I explore this form of rich spatiotemporally continuous active causal learning in the final empirical chapter.

2.8.5 Even more approximation

It is hard enough to do exact inference and intervention selection with CBNs. Once richer functional forms come into play the hypothesis, action and outcome spaces become even more formidable. Fortunately, richer data means the potential for stronger evidence about relationships, and principles of sample-based approximation still apply. Samples can be drawn from a prior and filtered, or the hypothesis space can be explored stochastically, as in MCMC sampling. While I will not model this process in as much detail for these richer active learning cases, Chapters 6, 7 and 8 will demonstrate that behavioural patterns are consistent with similar forms of approximation to those I model in the pure-contingency CBN cases I consider in the first half of the thesis.

2.9 Summary

In this chapter I set out a rational analysis of the related problems of causal structure learning, and intervention selection. I first developed the Causal Bayesian network framework, and the idea of learning through the contingency data produced by "variable fixing" interventions. I introduced a number of heuristic proposals and principled approximations that give insight into how bounded brains might approach this complex problem. I then generalised beyond the CBN case by introducing temporal and mechanistic statistics and correspondingly richer representations of causal structure. I provide the mathematical details of the CBN framework in a text box at the start of Chapter 3, for reference in the empirical chapters to follow, providing additional mathematical details relating to my treatment of time and physical causality in the last three chapters as and when they are needed.

Chapter 3

Learning causal structure through intervention

> "*The scientific mind does not so much provide the right answers as ask the right questions.*"
>
> — CLAUDE LEVI-STRAUSS

A number of studies have shown that people benefit from the ability to perform interventions during causal learning (Coenen et al., 2015; Lagnado & Sloman, 2002, 2004, 2006; Schulz, 2001; Sobel & Kushnir, 2006; Steyvers et al., 2003). However, only Steyvers et al. and Coenen et al.'s studies explored how people select what interventions to perform, and both only for the case of a single intervention on a single variable in a semi-deterministic context. In contrast, much real world causal learning is probabilistic and incremental, taking place gradually over many instances. It has not yet been explored in what ways sequential active causal learning might be shaped by cognitive constraints on memory and processing, or whether learners can plan ahead when choosing interventions.

Additionally, it has been shown that single-variable interventions are not sufficient to discriminate all possible causal structures (Eberhardt, Glymour, & Scheines, 2012). Interventions that simultaneously "control for" potential confounds to isolate a particular putative cause are cornerstones of scientific testing (Cartwright, 1989) and key to scientific thinking (Kuhn & Dean, 2005). However, the only interventional learning study that allowed participants to perform multi-hold interventions was Sobel and Kushnir (2006), and this study did not analyse whether participants used these interventions effectively.

A final point is that no previous studies have explicitly incentivised causal learners. There is ambiguity in any assessment of intervention choices based on comparison to a correct or optimal behaviour according to a single criterion value. This is because one cannot assume that the participants' goal was to maximize the quantity used to drive the analyses. In other areas of active learning research, researchers have run experiments to discriminate between potential objective functions that might underpin human active learning (e.g. Baron, Beattie, & Hershey, 1988; Gureckis & Markant, 2009; Meder & Nelson, 2012; Nelson, 2005; Nelson et al., 2010) but this is yet to be explored in the domain of causal learning.

Clearly, there are many aspects of active causal learning that call out for further exploration. Therefore, this chapter presents two experiments and modelling that extend the existing work along several dimensions. In particular, we explore:

1. Whether people can choose and learn from interventions effectively in a fully probabilistic, abstract and unconstrained environment.
2. To what extent people make effective use of complex "controlling" interventions as well as simple single variable fixes.
3. What objective function best explains participants' intervention choices: do they act to maximize their expected utility, probability of being correct, or to minimize their uncertainty?
4. Whether people choose interventions to learn in a step-wise, "greedy" way or whether there is evidence they can plan further ahead.
5. How people's causal beliefs evolve over a sequence of interventions. Is sequential causal learning biased by cognitive constraints such as forgetting or conservatism?
6. Whether people's interventions and causal judgments can be captured by simple heuristics.

The first two points can be addressed through standard analyses of participants' performances in various causal learning conditions. However, the latter questions lend themselves to more focused analyses of the dynamics of participants' intervention selections. Therefore, in the second half of the chapter we will explore these questions by fitting a range of intervention and causal-judgment models directly to the actions and structure judgments made by participants in the experiments. We compare different learning functions (*utility gain*, *probability gain* and *information gain*) and compare

greedy learning models to models that plan ahead; and assess the influence of potential cognitive constraints (forgetting and conservatism). We also explore the extent to which participants' behaviour can be similarly captured by simple heuristic models as by more computationally complex Bayesian models.

Throughout this thesis, but particularly in this chapter and Chapters 4 and 5, *causal Bayesian networks* (CBNs) are used as a framework for causal representation (Pearl, 2000; Spirtes et al., 1993). Thus, the adopted formalism is provided in the text box below for easy reference.

3.1 The Causal Bayesian network framework

3.1.1 Representation

In a CBN, nodes represent variables (i.e. the component parts of a causal system); arrows represent causal connections; and parameters encode the combined influence of parents (the source of an arrow) on children (the arrow's target). Following standard graph nomenclature, we refer to the space between a pair of nodes in a model as an "edge", so that an acyclic causal model defines each edge as taking one of three states: forward $\rightarrow$, backward $\leftarrow$, or unconnected $\varnothing$. Bayesian networks are defined by the Markov condition, which states that each node is independent of all of its non-descendants given its parents. Such graphs can represent continuous variables and any form of causal relationship; but here we focus on systems of binary $\{0 = \text{absent}, 1 = \text{present}\}$ variables and assume generative connections — meaning we assume that the presence of a cause will always *raise* the probability that the effect is also present. It is worth noting that these graphs cannot naturally represent cyclic or reciprocal relationships. However, there are various ways to extend the formalism as discussed later in the thesis.

To parametrise causal models, we assume Cheng's Power PC (1997) convention, which provides a simple way of capturing how probabilistic causal influences combine. This assumes that causes have independent chances of producing their effects, meaning the probability that a variable takes the value 1 is a noisy-OR combination of the *power* or strength w_S of any active causes of it in the model, together with that of an omnipresent background strength w_B encapsulating the influence of any

causes exogenous to the model (Glymour, 1998). We write $\mathbf{w} = \{w_S, w_B\}$. The probability that variable X_i takes the value 1 is thus

$$P(X_i = 1 | \, \mathrm{pa}(X_i), \mathbf{w}) = 1 - (1 - w_B)(1 - w_S)^{\sum_{X_j \in \mathrm{pa}(X_i)} X_j} \tag{3.1}$$

where $\mathrm{pa}(X_i)$ denotes the parents of variable x in the causal model (see Figure 3.1a for an example). For convenience, we will generally assume $\mathbf{w}$ is the same for all connections and components.[a]

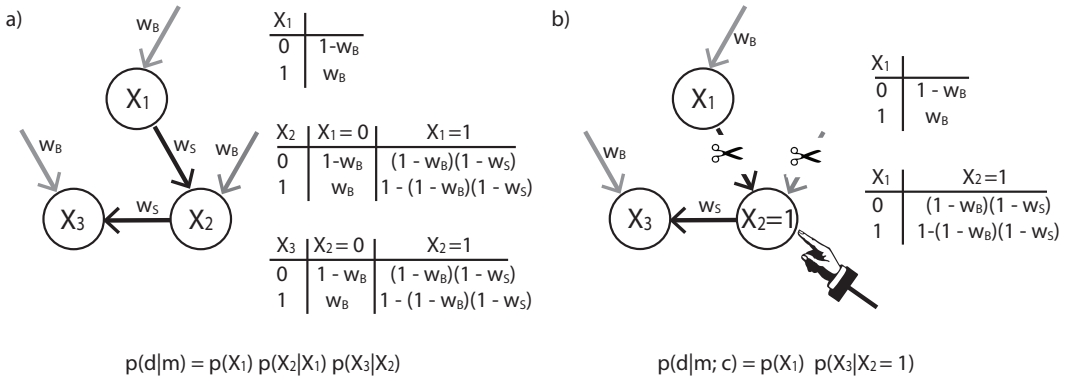


FIGURE 3.1: Causal model representation. a) An example CBN, parametrized with strength w_S and base rate w_B. The tables give the probability of each variable taking the value 1 conditional on its parents in the model and the omnipresent background noise rate w_B. b) Visualisation of intervention $\mathrm{Do}[X_2 = 1]$. Setting X_2 to 1 renders it independent of its normal causes as indicated by the scissors symbols.

3.1.2 Inference

Each causal model m over variables $\mathbf{X} = \{X_1, \ldots, X_n\}$ with strength and background parameters $\mathbf{w}$, assigns a probability to each datum $\mathbf{d} = \{X_1 = \mathrm{x}_1, \ldots, X_n = \mathrm{x}_n\}$, propagating information from the variables that are fixed through intervention $\mathbf{c}$, to the others (see Figure 3.1b). The space of all possible interventions $\mathcal{C}$ is made up of all possible combinations of fixed and unfixed variables, and for each intervention $\mathbf{c}$ the possible data $\mathcal{D}_\mathbf{c}$ is made up of all combinations of absent/present on the unfixed variables. We use Pearl's Do[.] operator (Pearl, 2000) to denote what is fixed on a given test. For instance, $\mathrm{Do}[X_1 = 1, X_2 = 0]$ means a variable X_1 has been fixed "on" and variable X_2 has been fixed "off", with all other variables free to vary.[b] Interventions allow a learner to override the normal flow of causal influence in a system, initiating activity at some components and blocking potential influences between others. This means they can provide information about the presence and

direction of influences between variables that is typically unavailable from purely observational data, without additional cues such as temporal information. For instance, in Figure 3.1.1b, we fix X_2 to 1 and leave X_1 and X_3 free ($\mathbf{c} = \mathrm{Do}[X_2{=}1]$). Under the $X_1 \rightarrow X_2 \rightarrow X_3$ model we would then expect X_1 to activate with probability w_B and X_3 with a probability of $1 - (1 - w_B)(1 - w_S)$.

In total, the probability of datum $\mathbf{d}$, given intervention $\mathbf{c}$, is just the product of the probability of each variable that was not intervened upon, given the states of its parents in the model

$$P(\mathbf{d}|m, \mathbf{w}, \mathbf{c}) = \prod\nolimits_{x \in (X \notin \mathbf{c})} P(x|\{\mathbf{d}, \mathbf{c}\}_{pa(x)}, \mathbf{w}), \tag{3.2}$$

where $\{\mathbf{d}, \mathbf{c}\}_{\mathrm{pa}(x)}$ indicates that those parents might either be observed (part of $\mathbf{d}$) or fixed by the intervention (part of $\mathbf{c}$).

In fully Bayesian inference, the true model is considered to be a random variable M. Our prior belief $P(M)$ is then an assignment of probabilities, adding up to 1 across possible models $m \in \mathcal{M}$ where $\mathcal{M}$ is the set of all possible models. When we observe some data $D = \{\mathbf{d}^i\}$, associated with interventions $C = \{\mathbf{c}^i\}$, we can update these beliefs with Bayes theorem by multiplying our prior by the probability of the observed data under each model and dividing by the weighted average probability of those data across all the possible models. We can condition on w_S and w_B if known

$$P(m|D, \mathbf{w}; C) = \frac{P(D|m, \mathbf{w}; C)P(m)}{\sum_{m' \in \mathcal{M}} P(D|m', \mathbf{w}; C)P(m')}, \tag{3.3}$$

or else marginalise over their possible values (see Appendix A).

We will typically treat the data as being independent and identically distributed, so $P(D|m, \mathbf{w}; C) = \prod_i P(\mathbf{d}^i|m, \mathbf{w}; \mathbf{c}^i)$.

If the data arrive sequentially (as $D^t = \{\mathbf{d}^1, \ldots, \mathbf{d}^t\}$; and similarly for the interventions), we can either store them and update at the end, or update our beliefs sequentially, taking the posterior $P(M|D^{t-1}, \mathbf{w}; C^{t-1})$ at timestep $t-1$ as the new "prior" $P^t(M)$ for datum $\mathbf{d}^t$ (or $P^t(M, \mathbf{w})$ if parameters $\mathbf{w}$ are also unknown).

3.1.3 Choosing interventions

It is clear that different interventions yield different outcomes, which in turn have different probabilities under different models. This means that which interventions are valuable for identifying the true model depends strongly on the hypothesis space and prior. For instance fixing X_2 to 1 ($\mathrm{Do}[X_2 = 1]$) is (probabilistically) diagnostic if you are primarily unsure whether X_1 causes X_3 because $p(X_3|\,\mathrm{Do}[X_2 = 1])$ differs depending whether $\mathrm{pa}(X_3)$ includes X_1 (see Figure 3.1.1b). However, it is not diagnostic if you are primarily unsure whether X_1 causes X_2 because X_2 will take the value 1 regardless of whether $\mathrm{pa}(X_1)$ includes X_2.

The value of an intervention can be quantified relative to a notion of uncertainty. We can define the value of an intervention as the expected reduction in uncertainty about the true model after seeing its outcome.[c] To calculate this expectation, we must average, prospectively, over the different possible outcomes $\mathbf{d}' \in \mathcal{D}_{\mathbf{c}}$ (recalling $\mathcal{D}_{\mathbf{c}}$ is the space of possible outcomes of intervention $\mathbf{c}$) weighted by their marginal likelihoods under the prior. For a greedily optimal sequence of interventions $\mathbf{c}^1, \ldots, \mathbf{c}^t$, we take $P(M|D^{t-1}, \mathbf{w}; C^{t-1})$ as our prior each time. The most valuable intervention $\mathbf{c}^t$ at a given time point is then

$$\arg\max_{\mathbf{c}\in\mathcal{C}} \mathop{\mathbb{E}}_{\mathbf{d}'\in\mathcal{D}_{\mathbf{c}}} \left[V(M|\mathbf{d}', D^{t-1}, \mathbf{w}; C^{t-1}, \mathbf{c})\right], \tag{3.4}$$

where $\mathbb{E}[.]_{\mathbf{d}'\in\mathcal{D}_{\mathbf{c}}}$ denotes the average over outcomes $\mathbf{d}'$ and $V(.)$ denotes the learner's objective function. The corresponding form for the case of unknown parameters $\mathbf{w}$ is also given in Appendix A.

We can generalise this greedy strategy to the case of an arbitrary prior belief $P^t(M)$, where the expected value of a given intervention is

$$\mathop{\mathbb{E}}_{\mathbf{d}'\in\mathcal{D}_{\mathbf{c}}} \left[V^t(M|\mathbf{d}', \mathbf{w}; \mathbf{c})\right]. \tag{3.5}$$

Objectives

Three commonly used objectives in the active learning literature are *(expected) utility gain* (Gureckis & Markant, 2009; Meder & Nelson, 2012), *probability gain* (Baron, 2005) and *information gain* (Shannon, 1951; Steyvers et al., 2003).

Utility gain

If you know how valuable correctly identifying all or part of the true causal system is, then the goal of your interventions is to get you to a state of knowledge about the true graph that is worth more to you than the one you were in before. Mathematically, this means maximisation of expected post-outcome, post-classification expected utility $\mathbb{E}_{\mathbf{d}' \in \mathcal{D}_\mathbf{c}} \left[U^t(M|\mathbf{d}', \mathbf{w}; \mathbf{c}) \right]$.

If each potential structure judgment b has a utility given that the true graph is m', we can capture the value of any judgment by some reward function R. Assuming one will always choose the causal structure with the highest expected reward, the *utility gain* $\mathrm{Ug}(M)$ of an intervention's outcome is the maximum over expected utilities of the possible judgments given the posterior $P^t(m|\mathbf{d}; \mathbf{c})$ minus the maximum for the prior $P^t(m)$:

$$\overset{t}{\mathrm{Ug}}(M|\mathbf{d}; \mathbf{c}) = \max_{b \in M} \sum_{m \in M} R(b, m) P^t(m|\mathbf{d}; \mathbf{c}) - \max_{b \in M} \sum_{m \in M} R(b, m) P^t(m). \tag{3.6}$$

An optimal intervention is defined as the intervention that maximizes the expected utility gain (i.e. replacing V by Ug in Equation 3.5).

Probability gain

While maximising expected utility can be seen as the ultimate goal of intervening, often a useful proxy is to maximize your expected probability of being correct. Under many normal circumstances choosing the most probable option will correspond to choosing the option that maximises your expected utility (Baron et al., 1988), however in terms of favouring one potential posterior distribution over another (as in planning interventions), the two values are more likely to differ depending on the reward function. Assuming you will choose the causal structure that is most probable, the *probability gain* $\mathrm{Pg}(M)$ can be written as:

$$\overset{t}{\mathrm{Pg}}(M|\mathbf{d}; \mathbf{c}) = \max_{m \in M} P^t(m|\mathbf{d}; \mathbf{c}) - \max_{m \in M} P^t(m). \tag{3.7}$$

An optimal intervention is defined as the intervention that maximizes the expected probability gain (i.e. replacing V by Pg in Equation 3.5)

Information gain

Another possible option for evaluating interventions comes from information entropy measures, which provide a way of measuring the overall uncertainty implied by a probability distribution. While there are a range of entropy measures (Nielsen & Nock, 2011), the most widely used is Shannon entropy (Shannon, 1951) $H(M)$, given by

$$H(M) = -\sum_{m_i \in M} P(m_i) \log_2 P(m_i). \tag{3.8}$$

Shannon entropy is largest for a uniform distribution and drops toward zero as that distribution becomes more peaked. We can call reduction in Shannon entropy *information gain* (Lindley, 1956) and use this as a way to measure the extent to which a posterior implies a greater degree of certainty across all hypotheses, rather than just improvement in one's post-decision utility or probability of making a correct classification. Information gain $\mathrm{Ig}(M)$ is given by

$$\overset{t}{\mathrm{Ig}}(M|\mathbf{d};\mathbf{c}) = \left[-\sum_{m \in M} P^t(m) \log_2 P^t(m)\right] - \left[-\sum_{m \in M} P^t(m|\mathbf{d};\mathbf{c}) \log_2 P^t(m|\mathbf{d};\mathbf{c})\right]. \tag{3.9}$$

An information gain optimal intervention is defined as the intervention that maximizes the *expected information gain* (i.e. replacing V by Ig in Equation 3.5)

$$\overset{t}{\mathrm{Eig}}(\mathbf{c}|M) = \mathop{\mathbb{E}}_{\mathbf{d}' \in \mathcal{D}_{\mathbf{c}}} \left[\overset{t}{\mathrm{Ig}}(M|\mathbf{d}', \mathbf{w};\mathbf{c})\right]. \tag{3.10}$$

[a] I also restrict myself to cases without any latent variable, although we note that imputing the presence of hidden variables is another important and computationally challenging component of causal inference (Buchanan, Tenenbaum, & Sobel, 2010; Kushnir, Gopnik, Lucas, & Schulz, 2010).

[b] We include the pure observation $Do[\varnothing]$ in $\mathcal{C}$.

[c] Strictly, this is greedy rather than optimal strategy because planning several steps ahead can result in a different intervention being favoured. The optimal choice, planning multiple steps steps ahead can be computed through dynamic programming (Puterman, 2009).

3.2 Comparing intervention objectives

Above, we introduced three commonly used objectives for driving active learning: *expected utility gain* (Ug, Equation 3.6), *probability gain* (Pg, Equation 3.7) and *information gain* (Ig, Equation 3.9). The extent to which these measures predict different intervention choices is one topic of investigation in this chapter. However, as a starting point we can consider what types of posterior distribution are high in utility, probability and information. To illustrate these differences, Figure 3.2 gives an example of three posterior distributions about which the three measures disagree. In the tasks we investigate here, people are rewarded according to how accurate their causal judgment is (e.g., how many of the causal connections and absences they correctly identify). This means that, according to expected utility, being nearly right is better than being completely wrong. Accordingly, we can expect that *utility gain* will prioritise interventions that divide the space of likely models into subsets of similar models rather than subsets of more diverse ones so that one is left with probable options that are all relatively rewarding. For example, in Figure 3.2, we see that a utility driven learner would want to prioritise discriminating m_3 from m_2 and m_1, because, whatever the outcome, they stand to make the same or more points if asked immediately after. If their intervention leaves them with m_1 and m_2 as candidates, they will still make 2.5 points on average by guessing between them (corresponding to a position halfway along the left face of the ternary plot), while if the outcome favours m_3 they can already make 3 points (corresponding to the bottom right corner of the ternary plot). Probability gain is only concerned with interventions likely to raise the probability of the most likely hypothesis, and does not care about similarity or overlap between hypotheses, or whether uncertainty between the various less-probable options is reduced. We see this in the ternary plot with the value of a location depending purely on its distance from one of the corners. Thus, we expect probability gain to favour interventions that are targeted toward confirming or dis-confirming the current leading hypothesis. In contrast, *information gain* concerns the reduction in uncertainty over all hypotheses. It will favour interventions that are expected to make a large difference to the spread of probability across the less probable networks, even when this will not pay off immediately for the learner in terms of increasing utility or probability of a correct classification. We see this in the curved contours in the ternary plot, showing a preference for distributions that are not only close to a corner but also to one side, indicating a second-place preference.

In support of the idea that *probability gain* might drive human active information search,

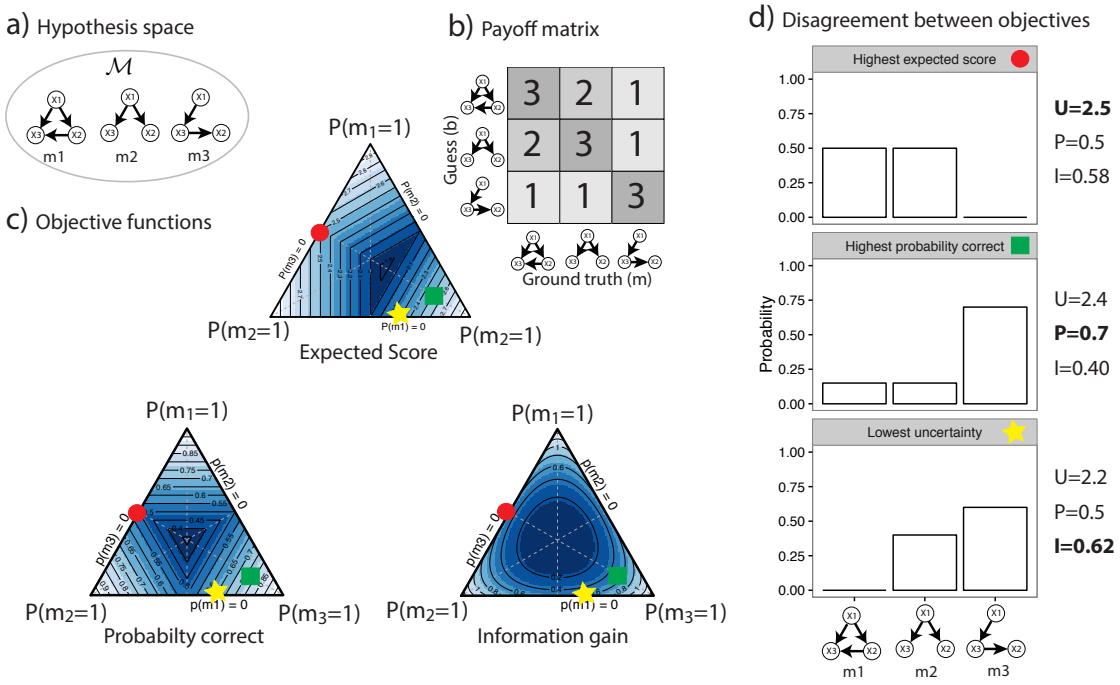


FIGURE 3.2: An example of differences in evaluation of posterior distributions with *expected (U)tility*, *(P)robability correct* and *(I)nformation*. a) A hypothesis space of three possible models. b) Payoff matrix assuming the learner is paid one point per correctly identified connection. c) Ternary plot visualisation of $V(M)$ for different objectives. d) An example three-way disagreement. Utility gain favours the top option, Probability the middle and Information the bottom.

Nelson (2010) has found participants' queries in a one-shot active classification task to be a closer match to *probability gain* than *information gain*. On the other hand, Baron, Beattie and Hershey's (1988) studies suggest that people will often select the question which has the higher *information gain* even if, for all possible answers, it will not change their resulting decision. There is also some recent evidence that people pick queries that are efficient in terms of *information gain* rather than *probability gain* in other areas of active learning (Gureckis & Markant, 2009; Meier & Blair, 2012). Steyvers et al (2003) used *information gain* to quantify the intervention chosen by participants in their task, but they did not compare this with other measures. For these reasons, when analysing our tasks we will consider *utility gain*, *probability gain* and *information gain* alongside one another, asking to what extent the measures imply distinct patterns of interventions, and to what extent people's active causal learning choices appear to be driven by one or other measure.

Greedy, or global optimisation?

Another issue here is that, when learning continues over multiple instances, greedily choosing interventions which are expected to obtain the best results at the next time point (whether in terms of information, highest posterior probability, or expected utility) is not guaranteed to be optimal in the long run. There may be interventions which are not expected to give good results immediately, but which provide the best results later on when paired with other interventions. To be truly optimal, a learner should treat intervention planning as a Markov decision process (Puterman, 2009), and look many steps ahead, always selecting the intervention which is the first step in the sequence of interventions which leads to the greatest expected final or total utility (assuming they will maximize on all future interventions). However, computing expectancies over multiple hypotheses and interventions when each intervention has many possible outcomes is computationally intractable (Hyafil & Rivest, 1976) for all but the smallest number of variables and most constrained hypothesis spaces, prohibiting optimal computation in the general case. It is an open question, which we will explore here, whether people can think more than one step ahead when planning interventions.

3.3 Experiment 1: Learning structure by intervention

In this first experiment, we test people's ability to learn causal structure through intervention. To do this we designed an interactive computer-based active learning task in Flash (see `http://www.ucl.ac.uk/lagnado-lab/el/nbt` for a demo). In the task, participants had to use interventions to find and mark the causal connections in several probabilistic causal systems.

Participants

Seventy-nine adults were recruited from Mechanical Turk for Experiment 1.[1] They were paid between $1 and $4 (M=$2.80) depending on performance.[2]

[1] Mechanical Turk (`http://www.mturk.com/`) is a web based platform for crowd-sourcing short tasks widely used in psychology research. It offers a well validated subject pool, diverse in age and background, suitable for high-level cognition tasks (Buhrmester, Kwang, & Gosling, 2011; Crump, McDonnell, & Gureckis, 2013; Gosling, Vazire, Srivastava, & John, 2004; Hauser & Schwarz, 2015; Mason & Suri, 2012).

[2] The number of participants was determined by our experimental budget of £200. Unfortunately participants' ages and genders were not stored in this Experiment.

Design and Procedure

For each problem, participants were faced with three filled grey circles, set against a white background. They were trained that these were nodes, and that they made up a causal system of binary variables but were not given any further cover story. Initially, all of the nodes were inactive, but when participants performed a test then some or all of the nodes could temporarily activate. An active node glowed green and wobbled from side to side, while an inactive one remained grey. For each structure participants would perform multiple tests before endorsing a causal structure and moving on to the next problem. The running score, test number and problem number were displayed across the top of the screen during testing. The locations of the three nodes (hereafter X_1, X_2 and X_3) were randomized and the nodes were not labelled.

Participants completed one practice problem and five test problems. The practice problem was randomly chosen from the five test problems. The test problems were presented once each, in a randomized order. Participants performed 12 tests per problem as described below before finalising their structure judgment and receiving a score.

Each test had three main stages (Figure 3.3):

1. First participants would select what intervention to perform. They could fix between 0 and 3 of the nodes either to *active* or *inactive*. Clicking once on a node fixed it to active (depicted with a '+' symbol), clicking again fixed it to inactive (depicted with a '−' symbol). Clicking a third time unfixed the node again. A pointing hand appeared next to fixed nodes to make it clear that they had been fixed by the participant.

2. Once the participant was happy with the intervention they had selected, they would press "Test" and observe the outcome of their test. The outcome would consist of 0-3 of the nodes activating. Whether a node activated on a given trial depended on the hidden causal connections and the choice of intervention. Participants were trained that nodes activated by themselves with a probability of .1 (unless they had been fixed, in which case they would always take the state they had been fixed in). They were also trained that causal links worked 80% of the time.[3] Therefore, fixing a node to active tended to cause any children of that node to activate and this would tend to propagate to (unfixed) descendants. The noise in the system

[3]Concretely, they had a causal power of .8. Combining causal power with the spontaneous activation rate, a node with one active cause had a $1 - (1 - .1)(1 - .8) = .82$ probability of activating.

meant that sometimes there were false positives where nodes activated without being caused by any of the other nodes, and false negatives where causal links failed to work. The pattern of data seen by a participant over the task was thus a partly random function of their intervention choices.

3. After each test there was a drawing phase in which participants registered their best guess thus far as to the causal connections between the nodes. Initially there was a question mark between each pair of nodes indicating that no causal link had been marked there yet. Clicking on these question marks during the drawing phase would remove them and cycle through the options *no link, clockwise link, anti-clockwise link*, back to *no link*. The initial direction of each link (clockwise or anti-clockwise) was randomized. Participants were not forced to mark or update links until after the final test but invited to mark as they went along as a memory aid. This approach was used to avoid forcing participants to make specific judgments before they had seen enough information to make an informed judgment, but to maximize our record of their evolving judgment during the task.

4. Participants performed 12 tests on each problem. After their last test, they were prompted to finalise their choice for the causal structure, i.e. they had to choose *no link, clockwise link* or *anti-clockwise link* for all three pairs of nodes, leaving no question marks. Once they had done this they were given feedback as to the correct causal structure and received one point for each correctly identified link (Figure 3.3). There were three node-pairs per problem (X_1-X_2,X_1-X_3 and X_2-X_3) and three options (*no-link, clockwise link, anticlockwise link*) per node-pair. This means that chance level performance was 1 correct link per problem, or $\approx$5 points over the five problems, while an ideal learner could approach 15 points. At the end of the task, participants received \$1 plus 20c per correctly identified link leading to a maximum payment of \$4.

Before starting the practice round, participants completed a comprehensive and interactive instructions section designed to familiarize them with the spontaneous activation rate; the causal power of the nodes; the role of the different interventions and the aim of the task. To train participants on the causal power of these connections, we presented them with a page with five pairs of nodes. The left node of each pair was fixed on and it was revealed that there was a causal connection from each left node to each right hand (unfixed) node. Participants were made to test these networks at least 4 times finding that an average of 4/5 of the unfixed nodes would activate. The outcomes of their first

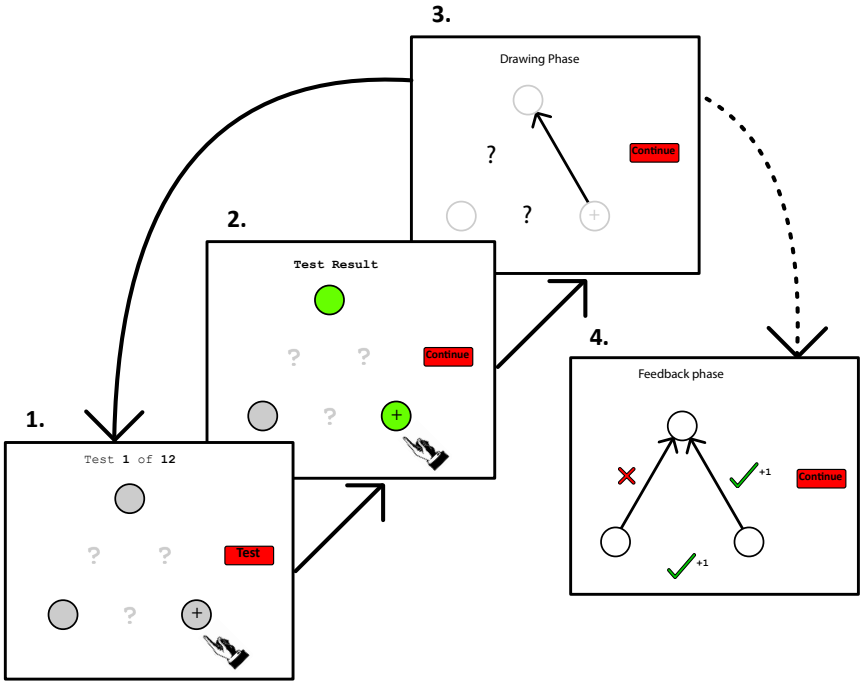


FIGURE 3.3: Experiments 1 and 2 procedure: 1. Choosing an intervention, 2. Observing the result. 3. Updating causal links. And, after 12 trials, 4. Getting feedback and a score for the chosen graph.

three tests were fixed to reflect this probability and thereafter the outcomes were generated probabilistically. Similarly, for the rate of spontaneous activations, participants were made to perform at least four tests on a page full of ten unfixed and unconnected nodes, where an average of 1/10 of these would activate on each test. In addition to this experience-based training, participants were told the probabilities explicitly. Before starting the task they had to answer four multiple choice questions checking they had understood: 1. The goal of the task (e.g., how to win money), 2. The role of fixing variables on and 3. fixing them off, and 4. The probabilistic nature of the networks. If the participant got less than 3 of 4 questions correct they were sent back to the beginning of the instructions.

Results

Participants identified an average of 9.0 out of 15 (SD = 4.1) causal links, and got 34% of the models completely right. This is well above the chance level of 5 out of 15 correct links (and 3.7% models correct), $t(78) = 8.60, p < .001$. However, the distribution of performance appears bimodal with one mode at chance and the other near ceiling (see Figure 3.4), suggesting that some participants were not able to solve the task while others

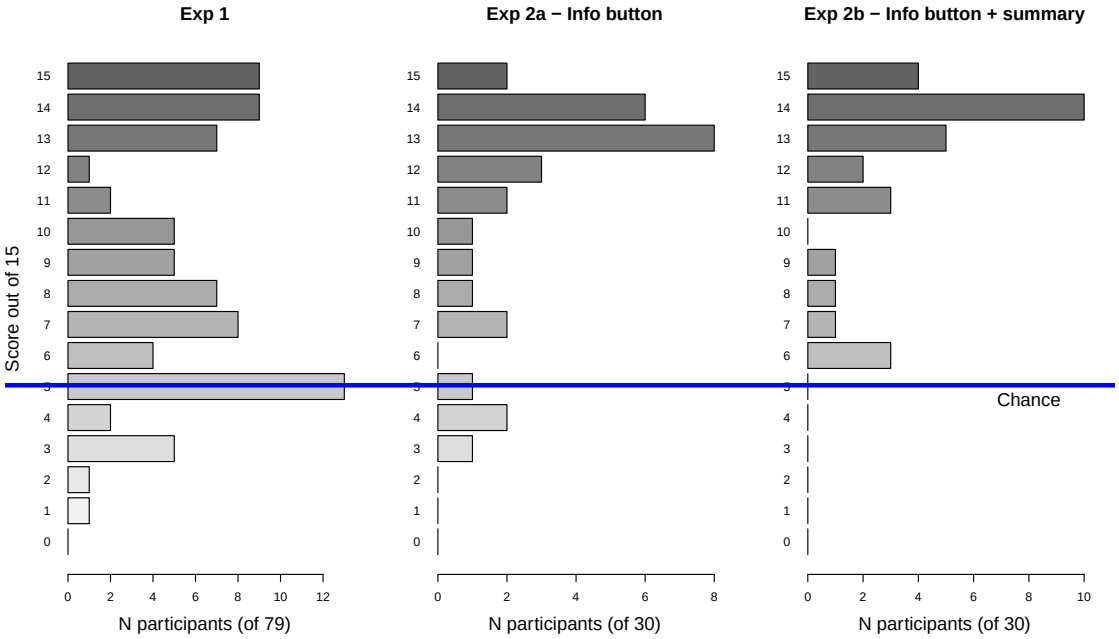


FIGURE 3.4: Histograms of scores in Experiment 1 and Experiment 2a and 2b. There were 15 points available in total (identifying all 15 connection-spaces correctly) and you could expect to get an average of 5 of these right by guessing (blue line).

TABLE 3.1: Three Most Frequent Judgment Errors in Experiments 1, 2a and 2b.

Exp	True structure	N correct	Mistaken for	N error
1	Chain $[X_1 \to X_2 \to X_3]$	20 (25%)	FC $[X_1 \to X_2 \to X_3, X_1 \to X_3]$	18 (23%)
1	Chain $[X_1 \to X_2 \to X_3]$	20 (25%)	Fork $[X_2 \leftarrow X_1 \to X_3]$	7 (9%)
1	FC $[X_1 \to X_2 \to X_3, X_1 \to X_3]$	26 (35%)	Chain $[X_1 \to X_2 \to X_3]$	7 (9%)
2a	Chain $[X_1 \to X_2 \to X_3]$	12 (40%)	FC $[X_1 \to X_2 \to X_3, X_1 \to X_3]$	12 (40%)
2a	Collider $[X_1 \to X_2 \leftarrow X_3]$	17 (56%)	FC $[X_3 \to X_1 \to X_2, X_3 \to X_2]$	3 (9%)
2a	Fork $[X_2 \leftarrow X_1 \to X_3]$	15 (50%)	FC $[X_3 \to X_2 \to X_1, X_3 \to X_1]$	3 (9%)
2b	Chain $[X_1 \to X_2 \to X_3]$	18 (60%)	FC $[X_1 \to X_2 \to X_3, X_1 \to X_3]$	8 (26%)
2b	FC $[X_1 \to X_2 \to X_3, X_1 \to X_3]$	17 (56%)	Chain $[X_1 \to X_2 \to X_3]$	4 (13%)
2b	Collider $[X_1 \to X_2 \leftarrow X_3]$	21 (70%)	FC $[X_3 \to X_1 \to X_2, X_3 \to X_2]$	3 (9%)

Note: "N correct" is the number of participants who identified this structure correctly and "N error" is the number of participants to make this particular error. FC is short for "Fully-connected"

did very well. This bimodality is confirmed by a dip test (Hartigan & Hartigan, 1985) $D = 0.09, p < .001$. There was no effect of problem order on performance $F(1, 394) = 0.06, \eta^2 = 0, p = 0.81$, nor did participants perform better on the problem they faced as their practice trial, and when they faced it again as a test problem $t(110) = -1.12, p = 0.26$. Participants did not over-connect or under-connect their final causal structures, on average opting for no-link for 30% of node-pairs, which was very close to the true percentage of 33%.

Participants were about equally accurate on the different structures, with slightly lower scores for the Chain, Fork and Fully-connected structures, than for the Collider or singly connected but there was no main effect of problem on score $F(4, 390) = 0.87, \eta^2 =$

$.01, p = 0.48$. However, looking at modal structure judgment errors, one error stands out dramatically: eighteen participants mistook the Chain $[X_1 \rightarrow X_2,\ X_2 \rightarrow X_3]$ for the $[X_1 \rightarrow X_2,\ X_1 \rightarrow X_3,\ X_2 \rightarrow X_3]$ Fully-connected structure, almost as many as participants who correctly identified the structure (Table 3.1).

3.4 Experiment 2: Information button and summary

Before analysing Experiment 1 further, an immediate question is why there was so much variance in participants' performance. One explanation for this could be that there are important individual differences between participants that strongly affected their ability to learn successfully. Steyvers et al.'s modelling suggested that people's ability to remember evidence from multiple past trials may be a critical psychological bottleneck for active causal learning. One way to check if poor performance stems from an inability to remember past tests is to provide participants with a history of their past interventions and their outcomes and assess whether this leads to better and more consistent causal learning. However, another perhaps simpler explanation for the variance is that some participants were confused about what to do and so responded randomly for all or much of the experiment.

To test both of these explanations, we ran another experiment using the same task as in Experiment 1 but with two additions. In Experiment 2a we provided an information button which would bring up a text box reminding them about what they were supposed to do at that stage of the task. In Experiment 2b, participants were still provided with this information button, but in addition they also were provided with a summary of all their past tests and their outcomes for the current problem. These were shown in a 4×3 grid to the left of the screen. After each test a new cell would be filled with a picture showing the causal system, the interventions selected (marked with "+" and "−" symbols as in the main task) and the nodes that activated (shown in green as in the main task, Figure 3.5).

Participants

Sixty additional Mechanical Turk participants aged 18 to 64 (M = 31.4, SD=11.2) completed Experiment 2. Once again, participants were paid between \$1 and \$4 (M=\$3.32, SD=.65).

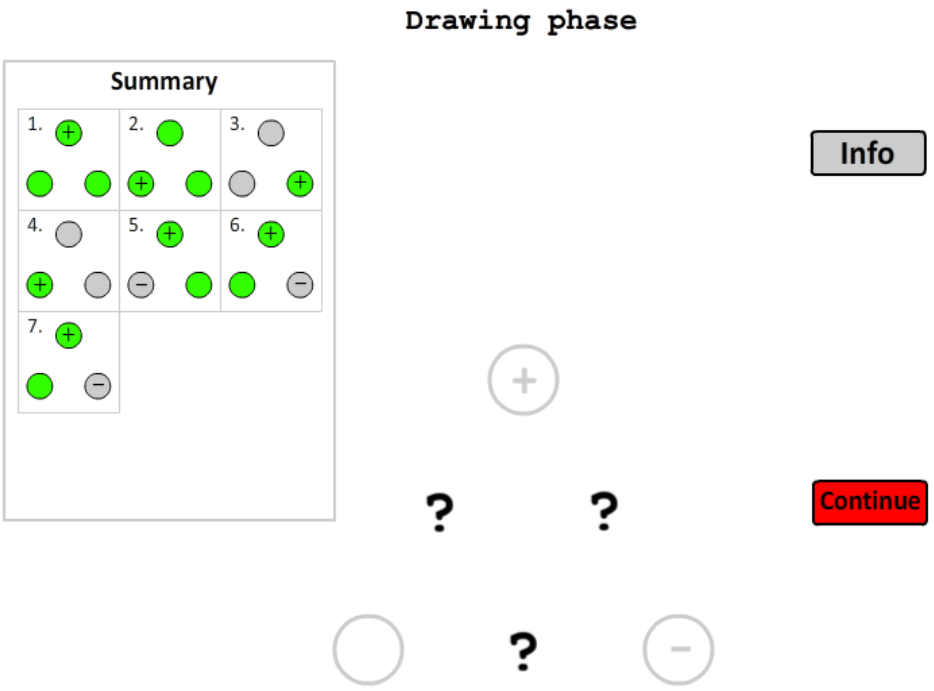


FIGURE 3.5: Interface in Experiment 2. Note the info button on the right (2a and 2b) and the summary information provided on the left (2b only: Nodes with a + symbol were fixed on, - symbol were fixed off, nodes with no symbol were unfixed. Green nodes activated and grey ones did not.)

Design and procedure

The procedure was exactly as in Experiment 1 except that now half of the participants were randomly assigned to Experiment 2a (info button only) and the other to Experiment 2b (info button + summary).

Results

On average, judgment accuracy in Experiment 2 was considerably higher than in Experiment 1 $t(138) = 4.2608, p < .001$. Participants in condition 2a (*info button only*) scored significantly higher at 11.1 (out of 15) correct links ($SD = 3.5$) than those in Experiment 1 $t(108) = 2.7$, $p = 0.009$ while participants in 2b (*info button + summary*) were slightly higher at 12.13 ($SD = 2.9$), again significantly higher than in Experiment 1 $t(108) = 4.5, p < .001$. However, the improvement from 2a to 2b was not significant $t(59) = 1.238, p = 0.22$. Inspecting Figure 3.4, we see that the number of participants performing close to chance is greatly reduced in both Experiment 2 conditions compared to Experiment 1 accounting for this difference in average performance.

These differences suggest that many of the poorer performers in Experiment 1 were simply confused about the task rather than being particularly poor at remembering evidence from past trials. However, scores were so high in Experiment 2 that failure to detect a performance level difference between conditions may be partly due to a ceiling effect. In line with this, we see that participants in the Experiment 2b *info + summary* were significantly faster at completing the task 18.4 ($SD = 8.1$) minutes than those in the Experiment 2a *info only* condition at 24.3 ($SD = 12.1$), $t(59) = 2.26, p = 0.03$. This suggests that the summary made a difference in terms of the effort or difficulty of at least some aspects of the task.

As with Experiment 1, there was no main effect of causal structure on performance in Experiment 2 $F(4, 295) = 0.64, \eta^2 = .008, p = 0.63$, nor were there any significant interactions between performance on the different structures and whether participants saw summary information (all p's$> .05$). However, as in Experiment 1, we found that participants were very likely to add a direct $X_1 \rightarrow X_3$ connection in for the Chain structure (Figure 3.8, Table 3.1).

We now move on to analyse the interventions selected by participants in the two experiments. Because the task in Experiments 1 and 2 was fundamentally the same, we will predominantly report analyses for all 139 participants together, but where relevant also explore differences between Experiment 1 and the two conditions of Experiment 2.

3.4.1 Intervention choices

Benchmarking the interventions

Participants' ultimate goal was to maximize their payout at the end of each problem (after their twelfth test). However, as mentioned in the introduction, there are various approaches to choosing interventions expected to help achieve this goal. Here we use three "greedy" (one-step ahead) value functions: expected *utility, probability* and *information gain* to assess how effectively participants selected different interventions.

To get a picture of the sequences of interventions favoured by efficient utility, probability or information seeking learners, we simulated the task 100 times using one-step ahead expected utility, probability and *information gain* (as defined in the introduction) to select each intervention. The prior at each time point was based on Bayesian updating from a flat prior using the outcomes of all previous interventions. All three measures always favoured "simple" interventions Do$[X_1 = 1]$, Do$[X_2 = 1]$ or Do$[X_3 = 1]$ (Figure 3.7)

for the first few tests for which the prior was relatively flat. Then, as they become more certain about the underlying structure, they increasingly selected "controlled" interventions with one node fixed on and another fixed off (e.g. $\text{Do}[X_1 = 1, X_2 = 0]$). After six tests, the probability that the models would select one of these controlled interventions was: .41 for the *probability gain* model, .37 for *probability gain* model and .51 for *information gain* model. For the later tests, if expected utility of the prior was already very close to 3 (full marks) and probability of correct classification was very close to 1, probability and *probability gain* were unable to distinguish between interventions, assigning them all expected gains of zero. Whenever this happened, these models would select interventions randomly. *Information gain* meanwhile continued to favour a mixture of simple, and controlled interventions. The *information gain* model would occasionally select an intervention with two nodes fixed on (4.5% of the time on tests 9 to 12). Other interventions (e.g., fixing two nodes off or fixing everything) did not provide any information about the causal structure so had expected gains of zero. These were only selected by the utility and *probability gain* models, and only on the last few trials when they could not distinguish between the interventions, so selected at random. The three approaches averaged scores of 14.1 (utility), 14.2 (probability), and 14.6 (information) correct links (Figure 3.6). Thus, within 12 trials it was possible for an efficient one-step ahead intervener to approach a perfect performance, averaging at least 14/15 depending on the choice of value function driving intervention choices.

Looking two steps ahead, the efficient active learners using one of these measures average almost identical average final scores (14.6, 14.4, and 14.6 points respectively) despite still using somewhat different sequences on interventions. The two-step-ahead models selected a higher proportion of controlled interventions than the greedy models (38% / 30% for *information gain*), and two-step ahead probability and *probability gain* were always able to distinguish between the interventions meaning they would no longer select interventions randomly on later tests.

For comparison, merely observing the system without fixing any variables would have provided very little information, capping a learner's ideal score at an average of 1.87 points per problem (9.35 overall, or a .26 probability of identifying the correct graph).

Participant's intervention choices

Efficiency of intervention sequences On average, participants selected highly efficient interventions in terms of utility, probability and information gain. Participants

Table 3.2: Comparison of the Proportion of Interventions of Different Types Selected by Participants and Simulated Learners.

	Proportion selected						
	Experiment			Simulation			
Intervention type	1	2a	2b	Random	UG	PG	IG
Observation	.01	.01	.02	.04	.08	.16	.00
Simple (e.g. Do$[X_1=1]$)	.73	.73	.77	.11	.34	.38	.68
Controlled (e.g. Do$[X_1=1, X_2=0]$)	.06	.09	.10	.22	.41	.30	.30
Strange 1 (e.g. Do$[X_1=1, X_2=1]$)	.11	.08	.05	.11	.07	.08	.02
Strange 2 (e.g. Do$[X_1=0]$)	.02	.02	.01	.11	.05	.05	.00
Strange 3 (e.g. Do$[X_1=0, X_2=0]$)	.00	.00	.00	.11	.05	.05	.00
Over-controlled (e.g. Do$[X_1=1, X_2=0, X_3=0]$)	.08	.07	.05	.30	.00	.00	.00

Note: Simulations selected interventions at (R)andom, and by maximising expected (U)tility, (P)robability, and (I)nformation gain.

learned much more than they would by picking interventions at random and they selected interventions that put their achievable score much closer on average to the benchmark models than to a random intervener in terms of their final expected utilities (see Figure 3.6).

Participants finished problems having learned enough that they could optimally score an average of 13.4 (SD=3.0) points per problem (M=13.1, SD=1.9 in Experiment 1, M=13.5, SD=.35 in Experiment 2) and have a .72 (SD=.25) probability of getting each graph completely right (M=.70, SD=.25 in Experiment 1, M=.75 SD=.23 in Experiment 2). This was significantly higher than selecting interventions at random, which would permit an average of only 11.6 points, or a .47 probability of getting each graph completely correct, $t(694) = 24, p < .001$. However, it was still significantly lower than what could be achieved by consistently intervening to maximize utility $t(694) = -12.7, p < .001$, probability $t(694) = -12.7, p < .001$ or information gain $t(694) = -18.3, p < .001$. The quality of participants' interventions was strongly positively associated with their ultimate performance. This is true for all measures of intervention quality tested here: *utility gain*: $F(1, 137) = 63, \eta^2 = 0.31, p < .001$, *probability gain*: $F(1, 137) = 81, \eta^2 = 0.37, p < .001$, and *information gain*: $F(1, 137) = 87, \eta^2 = 0.39, p < .001$.

Simple interventions As with the efficient learning models, "simple" interventions Do$[X_1 = 1]$, Do$[X_2 = 1]$ and Do$[X_3 = 1]$ were by far the most frequently selected, accounting for 74% of all interventions despite constituting only 3 of the 27 selectable interventions (Table 3.2). Propensity to use simple interventions was positively associated with performance across participants, $F(1, 137) = 41, \eta^2 = .23, p < .001$. As with the efficient learner models, the probability a participant would select a simple intervention

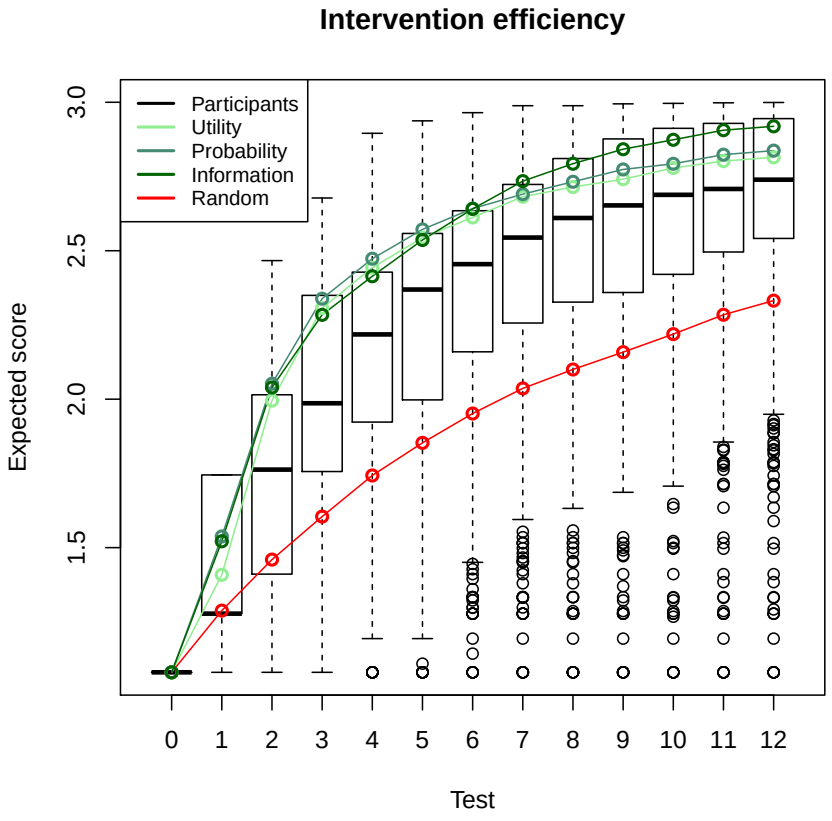


FIGURE 3.6: Participants' expected scores given the interventions they had selected thus far. Boxplot shows the best a participant could expect to score given the interventions and outcomes they had experienced up until that time point, averaged over the five problems. For comparison the other lines denote the mean expected scores of expected utility, probability or information maximising active learners (shades of green) and a passive learner who selects interventions at random (red), based on the simulations detailed in the text.

was highest at the start and then decreased over tests, $\beta - .03 \pm .007, Z = -4.1, p < .001$ (Figure 3.7).

Controlled interventions "Controlled" interventions (e.g. $\text{Do}[X_1 = 1, X_2 = 0]$) were selected only 7.4% of the time overall. This is not nearly as often they were selected by the efficient learner models. However, in line with these models, participants' probability of selecting a controlled intervention increased over tests, $\beta = .06 \pm .01, Z = 5.2, p < .001$. Propensity to use controlled interventions was also positively associated with performance, $F(1, 137) = 14.1, \eta^2 = .09, p < .001$. For each additional informative controlled intervention performed, participants scored .18 additional points in a task. The Chain and fully-connected structures are the two that cannot easily be distinguished without

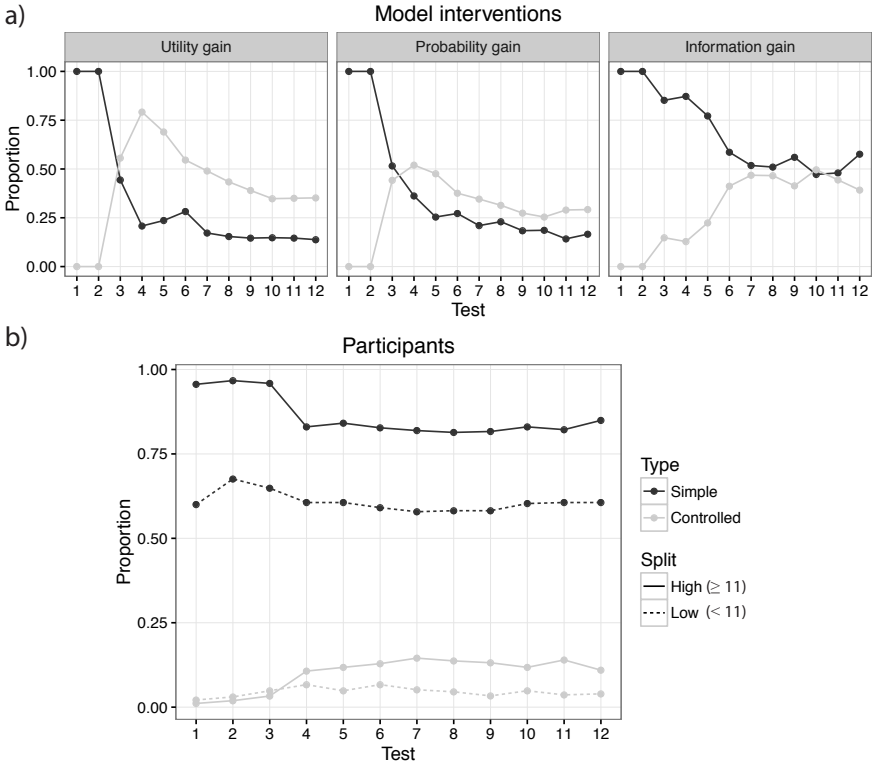


FIGURE 3.7: a) Proportion of "simple" (e.g. $\text{Do}[X_1 = 1]$) versus "controlled" ($\text{Do}[X_1 = 1, X_2 = 0]$) intervention choices for the three efficient learning models averaged over 100 simulations of the task. For later tests, based on increasingly peaked priors, expected *utility gain* and *probability gain* no longer distinguish between interventions and start to choose randomly while *information gain* continues to distinguish. b) Participants' proportion "simple" and "controlled" interventions over both experiments with a median split by performance.

a controlled intervention (see Figure 3.9), and accordingly we find use of controlled interventions is higher when the generating causal structure is a Chain or Fully-connected structure, $\beta = 0.50 \pm .08, Z = 6.0, p < .001$. In line with this, the use of controlled interventions also significantly predicts participants' probability of correctly omitting the $X_1 \rightarrow X_3$ connection in the Chain structure, $\beta = 1.7 \pm .4, Z = 4.6, p < .001$. In addition, a higher proportion of participants used controlled interventions at least once in Experiment 2 than in Experiment 1, $\chi^2(60) = 9.3, p = .002$ (41/60 compared to 44/79).

The fact that participants performed fewer "controlled" interventions in later tests than the benchmark efficient learner models is consistent with the idea that they were slower to learn. This would mean they would require more of the simple interventions to reach a level of certainty under which controlled interventions become the most valuable choice. The modelling in the next section will allow us to explore this possibility.

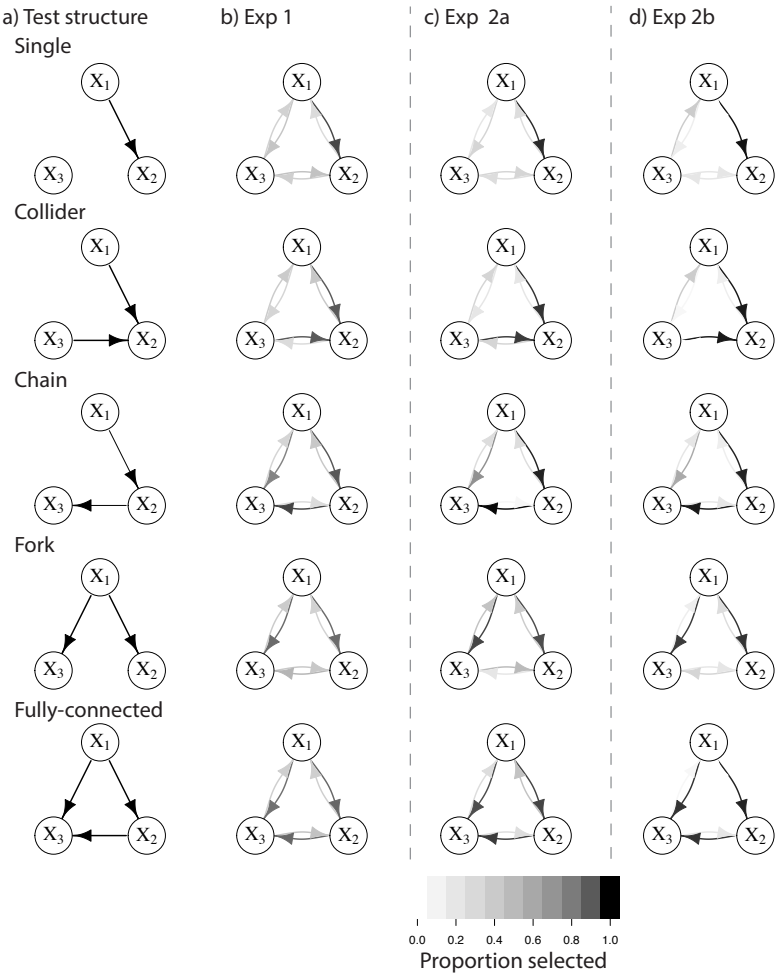


FIGURE 3.8: a) i.–v. Test structures used in Experiments 1 and 2. b)–d) Averaged final judgment for Experiments 1, 2a *info only* and 2b *info + summary*. Note that the direct link was often marked for the Chain (iii.).

Other interventions Participants sometimes selected interventions with two nodes fixed on (e.g. Do$[X_1 = 1, X_2 = 1]$), doing so 10% of the time. While the *information gain* model would select these interventions occasionally in later trials, participants were just as likely to select them early on $\beta = 0.009 \pm 0.01, Z = 0.77, p = .4$ and their propensity to select them was negatively associated with their performance, $F(1, 137) = 50, \eta^2 = .27, p < .001$. This suggests that participants typically did not use these interventions efficiently, or did not learn from them appropriately. Frequency of fixing everything (e.g. $X_1 = 1, X_2 = 0, X_3 = 0$) was strongly negatively correlated with performance, $F(1, 137) = 50, \eta^2 = .27, p < .001$. Participants who selected this type of intervention averaged final scores of only 8. Observing with no nodes fixed and fixing one or two nodes off were rarely selected and were not significantly associated with performance

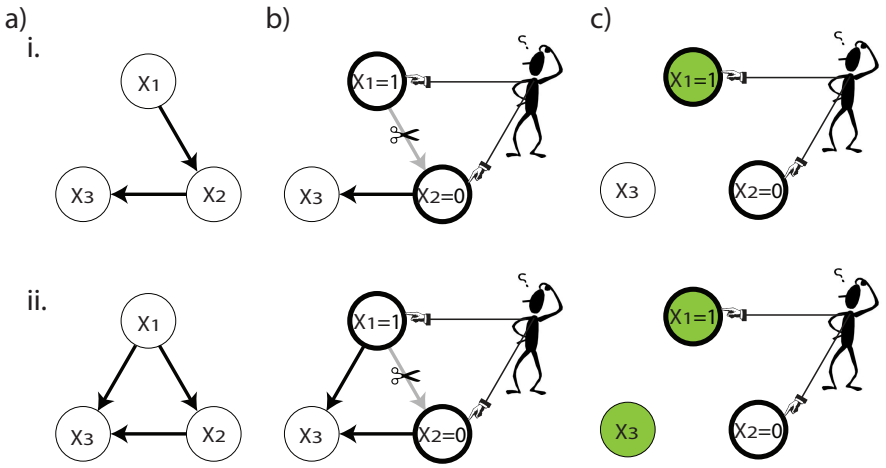


FIGURE 3.9: A controlled intervention. a) To distinguish i. $X_1 \rightarrow X_2 \rightarrow X_3$ from ii. $X_1 \rightarrow X_2 \rightarrow X_3, X_1 \rightarrow X_3$, one can manipulate X_1 while simultaneously holding X_2 constant. b) In the current context, this is achieved by fixing X_1 *on* and fixing X_2 *off*. c) If X_3 still turns *on* this is evidence for the ii., the fully-connected structure.

(p's of .14, .06 and .91 respectively).[4]

3.5 Modelling intervention selection and causal judgments

So far, we have analysed peoples' intervention selections at a relatively high level, looking only at how often particular types of intervention are chosen on average, either by good or bad participants, early or late during learning, or depending on the underlying causal structure. These high-level analyses have addressed the first two of our research questions, answering both in the positive:

1. The majority of people are able to choose informative interventions and learn causal structure effectively even when the environment is fully probabilistic, abstract and there is a large space of causal structures.

2. Most people can make use of complex "controlling" interventions to disambiguate between otherwise hard-to-distinguish structures. Ability to do this is a strong predictor of correctly identifying the causal structure, especially when the true structure is a Chain.

[4]Fixing two nodes off provides no information about the causal connections. Arguably, it still provides information about the spontaneous activation rate of variables but participants had already been trained on this in the instructions.

So far we have not touched upon the clear differences between interventions that fall within the same category (e.g. selecting $\mathrm{Do}[X_1 = 1]$ will provide very different information to $\mathrm{Do}[X_2 = 1]$ or $\mathrm{Do}[X_3 = 1]$). Additionally, we have not yet tried to distinguish which intervention selection measure is more closely in line with participants' choices. Looking across all three experiments, the three value functions favour different intervention(s) to one another on many of the participants' tests. *Utility* and *probability gain* disagree about what intervention should have been chosen on 19% of participants' tests. *Utility* and *information gain* disagree on 36% of participants' tests, and *probability gain* and *information gain* disagree on 39% of participants' tests. However, simply counting the frequency of agreement between participants' interventions and those considered most valuable by one or other measure is a blunt instrument for understanding participants' actions. The measures do not just give a single favoured intervention but a distinct value for each of the 27 possible interventions. Furthermore, the benchmark models assume perfect Bayesian updating after each intervention while a richer model comparison should allow us to compare the different measures while relaxing the assumption that participants are perfect Bayesians.

Thus, to progress further we will now fit and compare a range of models to participants' sequences of interventions and structure judgments. This will allow us to address our other research questions: 3. What objective function best explains people's choices, 4. Whether people can plan more than one intervention ahead, 5. Whether their belief update process is biased or constrained, and 6. Whether we can capture their active learning with simple heuristics.

On each test, a participant chooses an intervention but also can update their causal judgment by marking the presence or absence of possible causal links. The models discussed below will describe the intervention selections and causal judgments simultaneously, by assigning a probability to each intervention choice (from the 27 legal interventions) and to each combination of marked and unspecified links (out of the 27 possible combinations of causal connections). Free parameters are fitted to individual data since it is reasonable to assume that properties like memory and learning strategy are fairly stable within subjects but likely to differ between subjects in ways which may help us understand what drives the large differences in individual performance.

We fit a total of 21 models (Figure 3.10) separately to each participant's data. The models can be classified as either "expectancy-based" or "heuristic" models. The expectancy-based models assume that people choose interventions according to the expected value

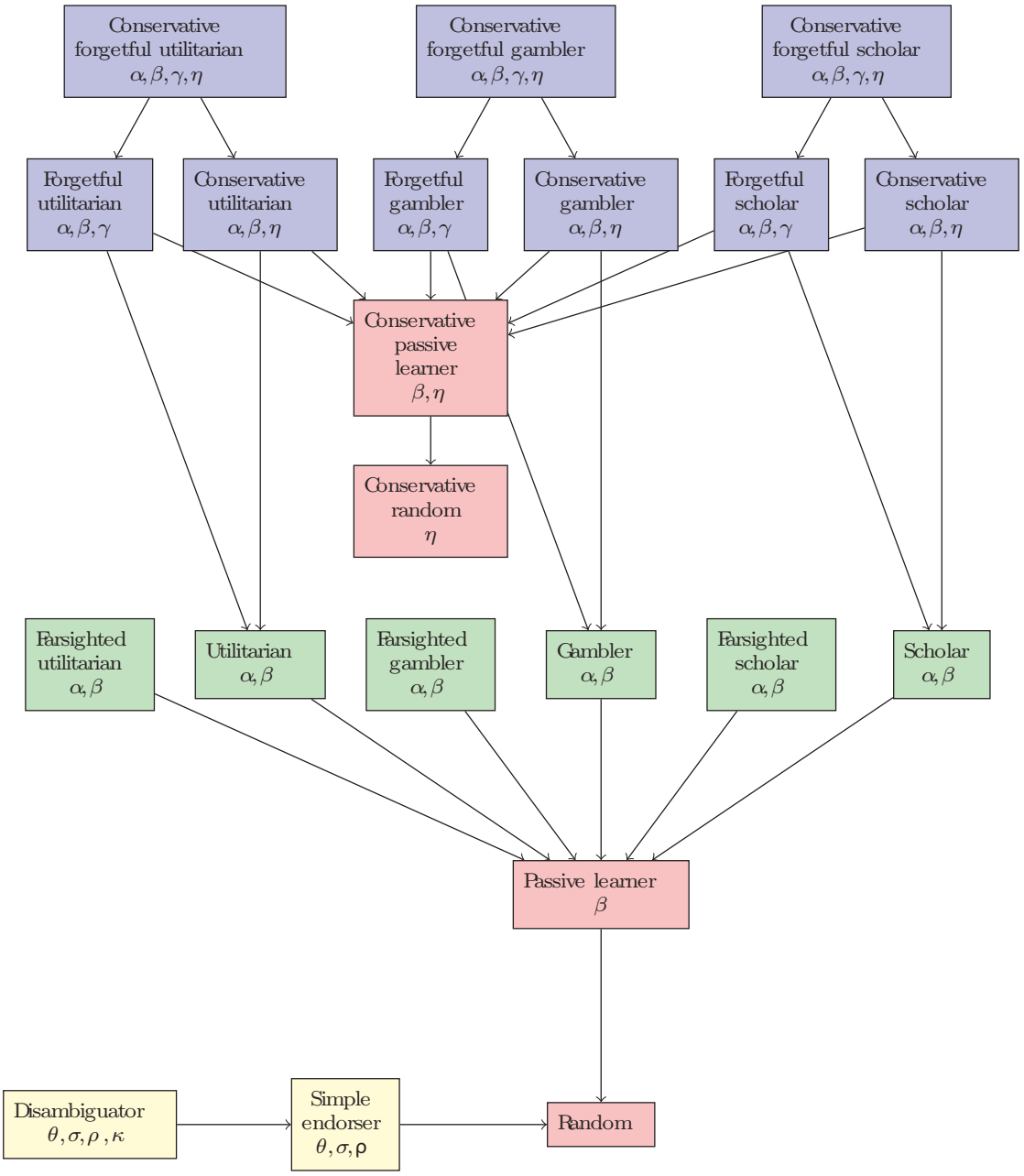


FIGURE 3.10: Schematic of the relationships between the models. Each model is nested within its parents and lists its fitted parameters. Blue rectangles indicate "bounded" models, green rectangles indicate "ideal" models, red rectangles indicate "null" models and yellow rectangles indicate "heuristic" models. Arcs representing the nesting of the conservative null models in the non-conservative null models are omitted for clarity

of each intervention, maximising either *utility*, *probability*, or *information gain* (see section 1.3 Quantifying Interventions). The models assume that the expectancies, as well as causal judgments, are based on Bayesian updating of probability distributions over the causal structures and the models are rational in the sense that they are optimal with respect to people's goals, although we also allow for the possibility of cognitive constraints such as forgetting and conservatism.

In Steyvers et al.'s (2003) study, many participants chose models that suggested they remembered only the result of their final intervention (having apparently forgotten or discounted the evidence from their previous observations) while others seemed to remember a little more. This is in line with what we know about the limited capacity of working memory (Cowan, 2001; Miller, 1956) and its close relationship with learning (Baddeley, 1992). Thus it seems likely that people are somewhat "forgetful", or exhibit recency with respect to integrating the evidence they have seen. We expect that in Experiment 2b, where a summary of past outcomes is provided, memory load should be reduced and participants should display less recency.

With regards to conservatism, research suggests that people interpret new data within their existing causal structure beliefs wherever possible (e.g. Krynski, Tenenbaum, et al., 2007). Anecdotally, people are typically slow or reluctant to change their causal beliefs. This suggests that people may also be *conservative* (Edwards, 1968) when updating their causal beliefs, even during learning. An additional motivation for this idea is the consideration that appropriate conservatism could actually complement forgetting; people may mitigate their forgetfulness about old evidence by remembering just what causal structural conclusions they have previously drawn from it (Harman, 1986). For example, suppose a participant registers an $X_1 \rightarrow X_2$ causal link after their first three interventions. We can take this as a (noisy) indication they are fairly confident at this stage that, whatever the full causal structure is, it is likely to be one with a link from X_1 to X_2. By the time the participant comes to their sixth intervention, they might not remember why they had concluded three trials earlier that there is an $X_1 \rightarrow X_2$ link, but they would still be sensible to assume that they had a good reason for doing so at the time. This means that it may be wise to be conservative, preferring to consider models consistent with links you have already marked, than those that are inconsistent even when you cannot remember why you marked the links in the first place.

In the heuristic models, intervention selections are not based on Bayesian belief updating and the expected value of interventions, but are derived from simple "rules-of-thumb". Although these models are not optimal with respect to any criterion, they can approximate the behaviour of the "rational" models reasonably well.

3.5.1 Expectancy-based models

We will call a model that assumes participants are pure pragmatists, choosing each intervention with the goal of increasing their expected score, a *utilitarian* model. A

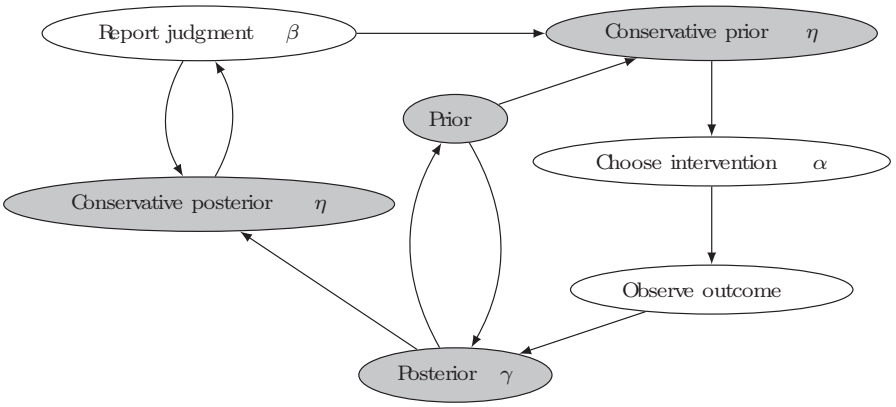


FIGURE 3.11: A flowchart of the expectancy based models. White nodes are observed quantities, grey nodes are unobserved quantities. Clockwise from the top right: The causal judgment reported at the previous time point and the prior distribution over causal structures combine to form a conservative prior. This is used to choose the next intervention. The outcome of the intervention is observed and this is integrated with the (partially forgotten) prior to arrive at the posterior distribution over causal structures. The posterior and the previously reported judgment are mixed to form a conservative posterior which influences the new judgment. The posterior becomes the new prior, then the process repeats.

utilitarian model assumes that participants choose interventions that are expected to maximize their payment at the next time point, or *utility gain* (see 3.1.3).[5] We will call a model that assumes people are just concerned with maximising their probability of being completely right (disregarding all other possible outcomes, or their payouts) a *gambler* model. A gambler model assumes participants choose actions which are expected to maximize the posterior probability of the most likely structure, or *probability gain*. We will call a model that assumes people try to minimize their uncertainty (without worrying about their probability of being right, or how much they will get paid) a *scholar* model. A scholar model assumes that participants choose actions to maximize their *expected information gain*, about the true structure at the next time point.

Updating causal beliefs and forgetting

All expectancy-based models assume that the learner's causal beliefs are represented by a probability distribution over all possible causal structures. At each time point, this probability distribution is based on Bayesian updating of their prior from the previous

[5]For each judgment the expected payout was calculated as the points received for that judgment summed over every possible graph, each multiplied by the posterior probability of the graph. As an example, endorsing Fork $[X_1 \rightarrow X_2; X_1 \rightarrow X_3]$ given the true structure is the Chain $[X_1 \rightarrow X_2; X_2 \rightarrow X_3]$ was worth 1 point because one of the three link-spaces (X_1-X_2) is correct while the other two are wrong.

time point to incorporate the evidence provided by the outcome of their latest intervention. However, rather than a complete Bayesian updating (Equation 3.3), we allow for the possibility that evidence from past trials may be partly discounted or forgotten.

There are various ways to model forgetting (Lewandowsky & Farrell, 2010; Wixted, 2004). A reasonable (high-level) approach is to assume that people will forget random aspects of the evidence they have received, leading to a net "flattening" of participants' subjective priors going into each new intervention. We can formalise this by altering the Bayesian update equation, such that a uniform distribution is mixed with the participants' prior on each update to an extent controlled by a forgetting parameter $\gamma \in [0, 1]$. So instead of

$$P^t(m|\mathbf{d}^t; \mathbf{c}^t) \propto P(\mathbf{d}^t|m; \mathbf{c}^t)P^t(m)$$

as in Equation 3.3, we have:

$$P^t(m|\mathbf{d}^t; \mathbf{c}^t) \propto P(\mathbf{d}^t|m; \mathbf{c}^t) \left[(1-\gamma)P^t(m) + \gamma\frac{1}{|m|}\right] \tag{3.11}$$

where $|m|$ is the total number of structures in $\mathcal{M}$, and distributions are computed recursively as $P^t(m) = P^{t-1}(m|\mathbf{d}^{t-1}; \mathbf{c}^{t-1})$. By setting γ to 0 we get a model with no forgetting and by setting it to 1 we get a model in which everything is forgotten after every test.

Choosing interventions

The expectancy-based models assume that intervention choices are based on the expected values of interventions. Let $v_{1t}, \ldots, v_{nt}$ denote the expected values $v_{\mathbf{c}t} = E_{\mathbf{d}}[V^t(M|\mathbf{c}, \mathbf{d})]$, where the generic function V is identical to Ug (Equation 3.6) in the utilitarian models, the *probability gain* (Equation 3.7) in the gambler models, and the *information gain* (Equation 3.9) in the scholar models. Note that these quantities are computed from the distributions $P^t(m|\mathbf{d}; \mathbf{c})$ and $P^t(m) = P^{t-1}(m|\mathbf{d}^{t-1}; \mathbf{c}^{t-1})$ as defined in Equation 3.11. We assume that chosen interventions are based upon these values through a variant of Luce's choice rule (Luce, 1959), such that the probability a learner selects intervention $\mathbf{c}$ at time t is given by:

$$P(\mathbf{c}^t) = \frac{e^{\alpha v_{\mathbf{c}t}}}{\sum_{k\in n} e^{\alpha v_{kt}}} \tag{3.12}$$

The parameter α controls how consistent the learner is in picking the intervention with the maximum expected value. As $\alpha \to \infty$ the probability that the learner picks the intervention with the highest expected value approaches 1 and the probability of picking

any other intervention drops toward 0. If $\alpha = 0$, then the learner picks any intervention with an equal probability, i.e. $P(\mathbf{c}^t) = \frac{1}{n}$, for all $\mathbf{c}^t \in C$.

Marking causal beliefs and conservatism

All expectancy-based models assume that learners' marked causal links are a noisy reflection of their current belief regarding the true causal models, as reflected by the posterior distribution $P^t(m|\mathbf{d}^t; \mathbf{c}^t)$. However, rather than using $P^t(m|\mathbf{d}^t; \mathbf{c}^t)$ directly, we allow for the possibility that the marking of causal beliefs may be subject to *conservatism*.

To allow for conservatism, we assume marked causal beliefs reflect a conservative probability distribution $P^{t*}(m|\mathbf{d}^t; \mathbf{c}^t)$, which is a distorted version of the current distribution $P^t(m|\mathbf{d}^t; \mathbf{c}^t)$ in which the probability of causal structures consistent with the already marked causal links is relatively increased. Technically, this is implemented by multiplying the probability of consistent causal graphs by a factor $\eta \in (0, \infty]$ and then renormalising the distribution.[6] The conservative probability distribution is given by:

$$P^{t*}(m|\mathbf{d}^t; \mathbf{c}^t) = \frac{\eta^{I(m)} P^t(m|\mathbf{d}^t; \mathbf{c}^t)}{\sum_{m' \in \mathcal{M}} \eta^{I(m)} P^t(m'|\mathbf{c}^t; \mathbf{d}^t)} \tag{3.13}$$

Where $I[m]$ is an indicator function with value 1 if the structure m is consistent with the currently marked links, and 0 otherwise. Marked links are assumed to be selected based on this conservative distribution. Then, this distribution is used to compute the values of the subsequent intervention options. For $\eta > 1$, sticking with already specified links is more likely than changing them all, other things being equal, while if $0 \leq \eta < 1$, this would lead to anti-conservatism. Unlike forgetting, which has an effect that accumulates over trials, the conservative distortion is applied "temporarily" on each trial when marking beliefs and choosing the next intervention, but discarded thereafter, such that the prior on trial $t + 1$ is $P_{t+1}(m) = P^t(m|\mathbf{d}^t; \mathbf{c}^t)$ and not $P^{t*}(m|\mathbf{d}^t; \mathbf{c}^t)$. By setting $\eta = 1$ we get a model which assumes participants are neither conservative nor anti-conservative.

[6] This parameter only does work once participants have registered their beliefs about at least some of the links, but this is the case on 91% of trials. On 76% of these trials participants had registered a belief for all three links, meaning that the conservativeness parameter up-weights the subjective probability of this one structure while they are selecting their new belief state and choosing the next intervention. This means that even if this learner's posterior is relatively flat due to forgetting, structures consistent with their marked links still stand out, leading them to behave as if they have selectively remembered information confirming these hypotheses. On the 24% of trials in which some but not all links remained unspecified, the conservativeness parameter led to the structures consistent with the established links being up-weighted, leading the learner to favour interventions likely to distinguish between these options - concretely there would be 9 structures consistent with one specified link, and 3 consistent with two specified links.

As for interventions, we assume that a learner marks causal links through a variant of Luce's choice rule. The marking of links on each trial was optional, and initially all links were unspecified. As a result, links were often left unspecified, in which case a set of models $\mathcal{S}$, rather than a single causal model, is consistent with the marked links.[7] To capture this, the models marginalize over all structures consistent with the links marked on a trial:

$$P(b^t) = \frac{\sum_{m \in \mathcal{S}} e^{\beta P^{t*}(m|\mathbf{d}^t;\mathbf{c}^t)}}{\sum_{m' \in \mathcal{M}} e^{\beta P^{t*}(m'|\mathbf{d}^t;\mathbf{c}^t)}} \tag{3.14}$$

where b^t is the stated belief after trial t. For example, if the participant has marked $X_1 \rightarrow X_2$, but has so far left $X_1 - X_3$ and $X_2 - X_3$ unspecified, then the model sums over the probabilities of all the graphs that are consistent with this link. If a participant has not marked any links then their belief state for that time point trivially has a probability of 1.[8] By setting β to zero we get a model which assumes that participants are unable to identify causal links above chance regardless of what evidence they have seen.

Null, ideal and bounded expectancy models

In summary, the expectancy-based models have four free parameters: α controls the degree to which the learner maximizes over the intervention values, β controls the degree to which the learner maximizes over their posterior with their link selections at each time point, γ controls the extent to which participants discount or forget about past evidence and η controls the extent to which participants are conservative about the causal links they mark. See Figure 3.11 for a flow chart of how the full expectancy based models work. By constraining the models such that combinations of these parameters are fixed, a nested set of expectancy-based models is obtained (Figure 3.10). Fixing parameters to a priori sensible values can be important. For instance, we can assess whether a learner is forgetful by comparing a model in which the γ parameter is estimated to one in which the parameter is fixed to $\gamma = 0$.

A useful way to break down these models is divide them into "null" models, "ideal" models and psychologically "bounded" models. We will call models with one or both of α and β fixed to zero "null" models. These models either assume that no active

[7] A side effect of this aspect of the design is that we have more data on some participants than others. Those who rarely marked links before the end of the task reveal less information about how their belief at one time point influences their belief at subsequent time points.

[8] Cyclic Bayesian networks cannot be defined within the Bayesian network framework and participants were instructed that they were impossible during the instructions. Therefore, on the 4.3% of trials in which participants marked a cyclic structure ($[X_1 \rightarrow X_2, X_2 \rightarrow X_3, X_3 \rightarrow X_1]$, or $[X_1 \rightarrow X_3, X_3 \rightarrow X_2, X_2 \rightarrow X_1]$) their belief state was treated as unspecified so that the model did not return a likelihood of zero for that participant.

intervention selection takes place ($\alpha = 0$, interventions are selected randomly) and/or that no successful causal learning takes place ($\beta = 0$). We will call the models in which γ is fixed to 0 and η to 1 "ideal" models. These models are ideal in the sense that they set aside potential psychological constraints and so are at the computational level according to Marr's hierarchy (Marr, 1982). Comparing just these models addresses the question of which computational level problem participants' actions and judgments suggest they are (approximately) solving. Finally, we will call the full models in which one or both of γ and η are fit to the data "bounded" models. These models are bounded in the sense that they attempt to capture how psychological processing constraints potentially distort or change the computational problem, allowing us to explore how people might mitigate this in their intervention strategies.

Sensible evaluation of the bounded models requires different null models. For example, it may often be the case that someone is conservative about their beliefs despite those beliefs being completely random ($\beta = 0$). Alternatively people might be conservative *passive* learners yet unable to select sensible interventions, choosing interventions which are not more useful than chance ($\alpha = 0$). In these cases, we would have no reason to ascribe scholarly, gamblerly or utilitarian behaviour despite our models capturing some systematicity in participants' data.

Far-sighted scholars, gamblers and utilitarians

As mentioned in *Greedy or global optimisation?*, the values of different actions depend to some extent on how far the learner looks into the future. Computing expected values looking more than two steps ahead quickly becomes intractable even in the three-variable case, but we were able to compute the "ideal" two-step-ahead models for the three measures.[9] This allows us to check if there is evidence that people are able to look more than one step ahead when choosing interventions. Accordingly, we fit additional *farsighted utilitarian, gambler* and *scholar* models in which the intervention values for looking one step ahead were replaced with those looking two steps ahead. We can compare these to the one-step ahead "ideal" models to see if there is evidence that participants were planning more than one step ahead. We did not include freely estimated forgetting (γ) or conservatism (η) parameters in these models because recomputing the two-step ahead intervention values on the fly for different γ and η increments was prohibitively computationally expensive.

[9]These expectancies are computed recursively, taking the maximum over the second set of interventions and passing these values back to the first set of interventions.

3.5.2 Heuristic models

In addition to the various expectancy based models described above, we explored whether people's intervention patterns can be well described by heuristic active learning models. By heuristic models, we mean models in which probabilities are not explicitly represented and values are not calculated for different interventions. Instead, these models assume that learners follow simple rules of thumb in order to choose their interventions, and update their causal models, without performing computationally demanding probabilistic information integration (Gigerenzer et al., 1999). Here we fit two models, the first nested in the second.

The simple endorser

One way to significantly simplify the causal learning problem is to ignore the dependencies between the causal connections in the possible graphs (Fernbach & Sloman, 2009). Thus, if intervention $\mathrm{Do}[X_1 = 1]$ is performed and both X_2 and X_3 activate this can be seen as evidence for an $X_1 \rightarrow X_2$ connection and, independently, evidence for an $X_1 \rightarrow X_3$ connection. In contrast a full Bayesian treatment would also raise the probability of other hypotheses (the Chains and Fully-connected structures). Another way to simplify the problem is to ignore the Bayesian accumulation of probabilistic evidence and rather update belief directly to be consistent with the latest evidence. Concretely, in this task these assumptions would lead to people simply fixing variables on, one at a time, and adding links to any other nodes that activate as a result (removing any links to other nodes which do not activate as a result). We can operationalise this with a three parameter model (Figure 3.12) which selects one of the simple interventions with probability $\theta \in [0, 1]$ or else selects anything else with probability $1 - \theta$. With a probability $\sigma \in [0, 1]$ the belief state is updated such that it becomes the prior belief state B plus links L from the current fixed node to any activated nodes (and minus those not in L but in B), while with probability $1 - \sigma$ it either: stays the same (with probability $\varrho \in [0, 1]$) or takes any other state (with probability $1 - \varrho$). A potential strength of this model in fitting the data is that it leads to systematic misattribution of a Fully-connected structure when the true structure is a Chain, a behaviour exhibited by many participants. This happens because when the true structure is a Chain, intervening on the root node will tend to lead to both other nodes activating, leading to the addition of direct links from the root node. When the middle node is intervened on this will tend to activate the last node, leading to the addition of the third link. To the extent that

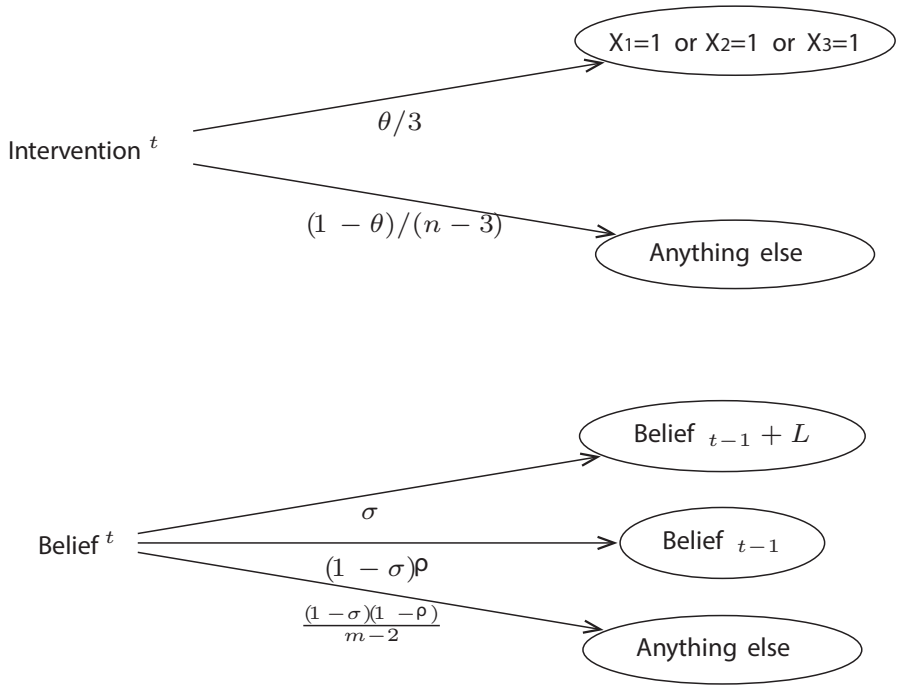


FIGURE 3.12: Process trees for the simple endorser. The learner follows an arrow with the probability written under the arrow and takes the action in the end node.

participants frequently act in this way, θ and σ will be high and the model fit will be good, and to the extent that they act in other ways the model will approach the fit of the null model in which beliefs and actions are selected at random.

The disambiguator

As we show earlier in the chapter, controlling variables is a hallmark of scientific thinking, and a necessary part of successfully disambiguating causal structures (Cartwright, 1989; Kuhn & Dean, 2005). In this task this takes the form of a controlled intervention in which one node is fixed on and another fixed off (e.g. $\text{Do}[X_1 = 1, X_2 = 0]$), normally performed after observing some confounding evidence (i.e. when you fix one node on and both other nodes activate). This action tests whether the node that remains unfixed is a direct effect of the node which is fixed on (Figure 3.9), and thus *disambiguates* between the structures which could explain why both unfixed nodes activated on the previous trial. In the general case, the putative cause node would remain fixed on, a single putative effect node would be left unfixed and the other $N - 2$ nodes would be fixed off.

The model is operationalised as selecting $\text{Do}[X_1 = 1]$, $\text{Do}[X_2 = 1]$ or $\text{Do}[X_3 = 1]$ or a disambiguation step (e.g. $\text{Do}[X_1 = 1, X_2 = 0]$, etc) with probability θ and something

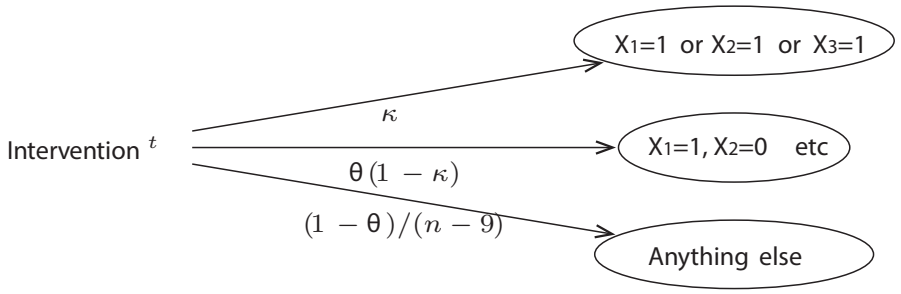


FIGURE 3.13: Process tree for the intervention selection step for the disambiguator. Belief update step is the same as for the simple endorser.

else with probability $1 - \theta$ (Figure 3.13). Propensity to select a simple endorsement step rather than a disambiguation step is governed by a fourth parameter $\kappa \in [0, 1]$. If a disambiguation step is performed and the unfixed node does not activate, then any connection from the activated node to the unfixed node is removed with probability σ. The belief update step is otherwise the same as for the simple endorser.

3.5.3 Model estimation and comparison

All models were fitted to individual's data by maximum likelihood estimation.[10] These consist of four nested sets, one for each of the three expectancy measures (*probability gain*, *probability gain* and *information gain*) and one for heuristic models. Each nested model has between zero and four parameters.

McFadden's pseudo-R^2 is computed for each model to give an idea of its goodness of fit.[11] This measure does not penalize model complexity so models are compared throughout using the Bayesian information criterion (BIC, Schwarz, 1978) which can be used to compare both nested and non-nested models (Lewandowsky & Farrell, 2010).

3.5.4 Model fit results

Full results of the model fits are contained in Table 3.3. Overall, the best fitting model was the fully bounded scholar model based on maximising *information gain* with both conservatism and forgetting (hereafter *CF scholar*). This model had a pseudo-R^2 of .47 indicating a very good fit to the data[12] and was the best fitting model for 103 out

[10] We used the Nelder-Mead algorithm to numerically maximize the likelihood, as implemented in R's `optim` function. Optimisation was validated by repetition with different starting parameter values.

[11] McFadden's pseudo-$R^2 = 1 - \frac{\log L(M_{\text{full}})}{\log L(M_{\text{minimal}})}$, where $L(M)$ denotes the likelihood of model M. The minimal model M_{minimal} is *random (no learning)* in Table 3.3, where both actions and endorsements are completely random.

[12] Values between .2 and .4 are considered a good fit (Dobson, 2010).

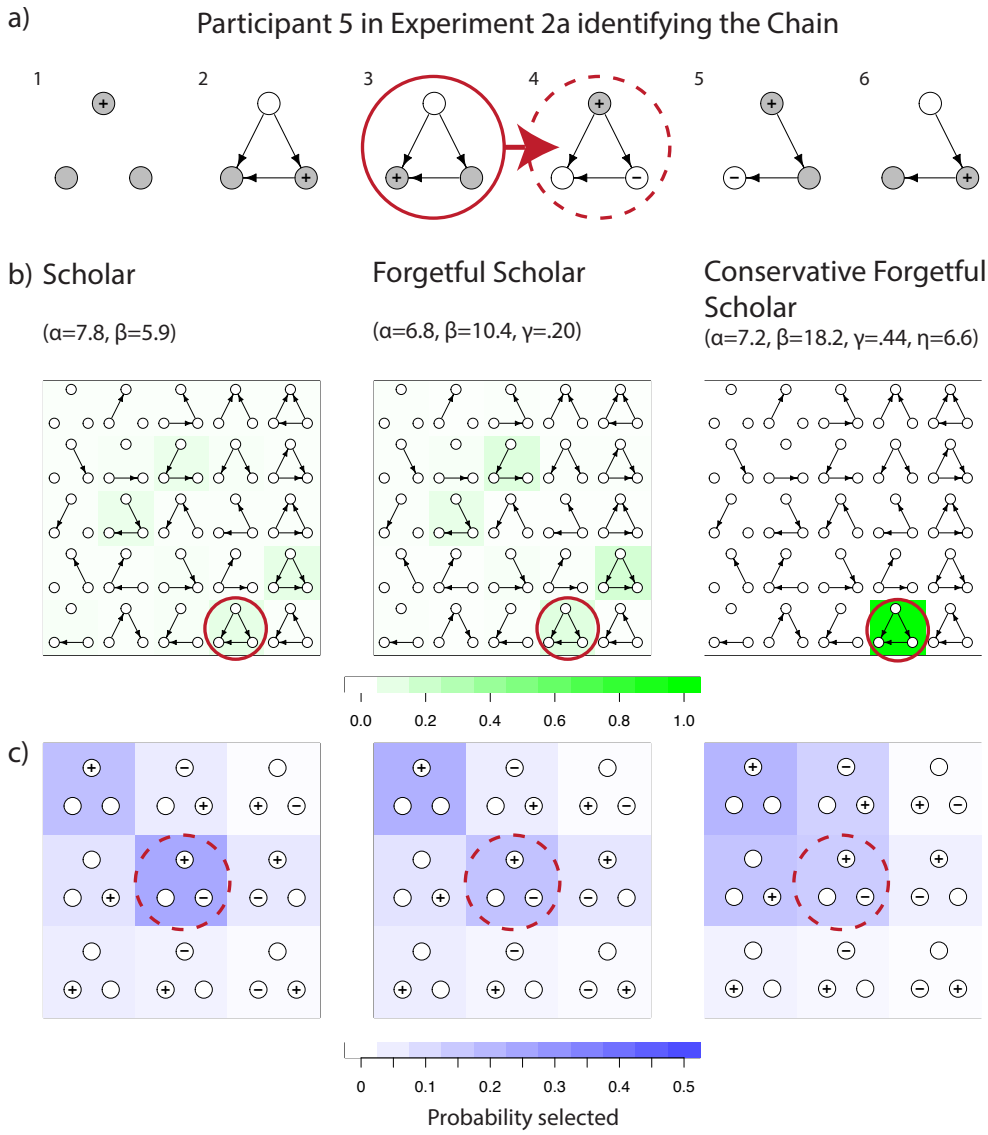


FIGURE 3.14: Visualisation of models. a) First 6 trials for participant 5 in Experiment 2a, identifying the Chain ($X_1 \rightarrow X_2, X_2 \rightarrow X_3$) structure. "+" and "−" symbols indicate interventions setting variables to 1 and 0. Grey nodes indicate the resultant activations, and the arrows replicate those marked by the participant at each time point. b) the probability that the participant registers each causal structure according to the *Scholar*, *Forgetful scholar* and *Conservative forgetful scholar* models (their actual choice is the full red circle) and c) the probability of selecting each of the "simple" and "controlled" interventions on the next test (actual choice is the dashed red circle).

of the 139 participants over Experiment 1, 2a and 2b according to the BIC. Of the 36 participants that were not best described as CF scholars, 24 were in Experiment 1 and many of these were best fit by the *conservative random* null model. See Figure 3.14 for a visual comparison of the scholar model with either or both of forgetting and conservatism as fit to a participant in Experiment 2a. Looking at their average scores, we see that those best described as *CF scholars* perform much better than those who are not *CF scholar* mean = 11.3, non-*CF scholar* mean = 6.6, $t(137) = 7.1, p < .001$.

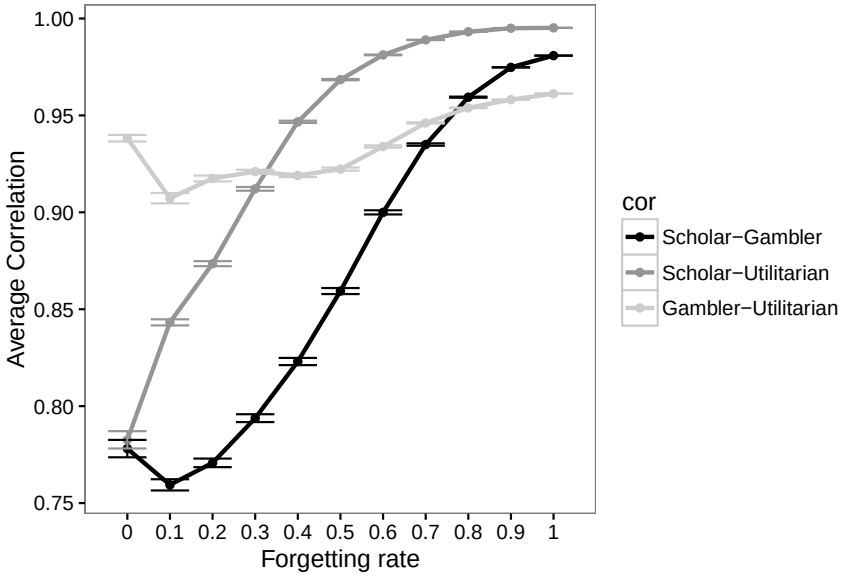


FIGURE 3.15: Mean correlation between intervention values according to these models for all participants' tests, as a function of forgetting.

Inspection of the *forgetful* models suggests that participants forgot a large amount of the evidence they received with median forgetting rates (γs) of .68, .79 and .47 for the *utilitarian, gambler* and *scholar* models respectively. When paired with conservatism in the conservative models, forgetting rates become even higher. This makes intuitive sense since high conservatism can result in a high probability for already marked links which would otherwise have to be due to participants maintaining more of the true posterior (see Figure 3.14). Looking at the parameter estimates of the *CF scholar* model, more forgetful people were also more conservative, with a significant rank-order correlation between γ and η ($\varrho = .43, p < .001$). In addition both forgetting and conservatism are negatively correlated with participants' overall scores, $\varrho = -.70, p < .001$ for γ and $\varrho = .53, p < .001$ for η.

Looking across experiments, we see that median forgetting (γ) in the *forgetful scholar* model drops considerably going from .71 in Experiment 1 to .30 in Experiment 2a and slightly further again to .25 in Experiment 2b. Naively we might expect that participants in Experiment 2b should not need a forgetting parameter, since they could see all of their past actions and outcomes. However, only one participant in Experiment 2b, and none in Experiment 2a or 1 was better fit by a model without a forgetting parameter meaning that the parameter still did work even for participants in Experiment 2b.[13] Rather we conclude that "forgetting" in our models does not just capture people's inability to recall past evidence. More generally, we think it captures a recency bias or tendency to

[13]This participant identified every connection correctly and was best described as an "ideal" scholar.

Table 3.3: Model Fitting Results.

Model	α	β	γ	log(η)	N best	N ideal	N heuristic	R^2	BIC
Random					0	5	1	0	98396
Passive		5			0	9	0	.15	84487
U	6.1	4.9			0	1		.20	79905
G	11	4.9			0	1		.19	80718
S	6	4.9			1	117		.25	75236
2-step U	9.3	4.9			0	0		.21	78576
2-step G	13	4.9			0	1		.21	78715
2-step S	4.1	4.9			0	5		.24	76229
C Random				4.5	9	0	28	.23	76574
C Passive		5.7		6.7	0	0	0	.25	75453
CU	7.1	5.8		3.1	0			.28	72635
CG	12	5.5		2.8	0			.27	73739
CS	6.7	5.8		3.3	0			.33	68190
FU	11	14	.68		4			.33	67694
FG	32	12	.79		1			.31	70027
FS	7.7	11	.47		2			.37	64039
CFU	15	16	.93	2	5			.43	58413
CFG	53	17	.97	1.3	6			.40	61680
CFS	**8.3**	**13**	**0.81**	**2.3**	**103**			**.47**	**54757**
	θ	ρ	σ	κ					
Simple	.85	.22	.63		5		61	.36	64985
Disambiguator	.95	.22	.63	.96	3		49	.37	64100

Note: Median parameter estimates, pseudo-R^2s and total BICs for all models. (C)onservative and/or (F)orgetful (U)ilitarian (based on Ug), (G)ambler (based on Pg) and (S)cholar (based on Ig) models . Additionally "N best" gives the number of participants best fit by each model according to BIC. "N ideal" gives the same statistics as "N best", but only includes the ideal learner models and appropriate null models, "N heuristic" does the same for the heuristic models.

attend disproportionately toward newer over older evidence regardless of whether the older evidence is still accessible.

"Ideal" models

Although the *CF scholar* model performed best overall, the *scholar, gambler* and *utilitarian* model predictions were often relatively similar when all four parameters were included. This could be because for flatter posteriors, the intervention values according to these models do not differ as much as they do when the posteriors are more peaked. Comparing predictions of the models with increasing forgetting rates confirms this (Figure 3.15), with the level of agreement about the best intervention(s) becoming close to 1 as forgetting rate increases toward 1. For a clearer assessment whether learners are best described as scholars, gamblers or utilitarians, we turn to a comparison of the "ideal" versions of these models (without forgetting or conservatism).

Considering only the "ideal" models and the relevant "null" models (Figure 3.10 and the "Best ideal" column in Table 3.3), the *scholar* model clearly outperforms the *utilitarian* and *gambler* models. In this set of models, the scholar best captures 117 out of the

139 participants, including almost all high scoring participants (*scholar* mean = 10.8, non-scholar mean 6.4). Nine of the poorest participants (average score 5.5) were also better described as achieving some learning, despite failing to select interventions more useful than chance (*passive learner*), but none were best fit by the completely chance level model *random* in which both α and β were set to zero.

Looking across experiments, we see that median αs for the ideal *scholar* model, controlling maximisation over intervention values, increase from 5.2 (SD=2.7) in Experiment 1 to 6.6 (SD=3.0) in Experiment 2a and 7.1 (SD=3.0) in Experiment 2b. Likewise, median βs, controlling maximisation over the posterior under the *scholar* model, increase from 3.5 (SD=26) in Experiment 1 to 5.9 (SD=3.5) in Experiment 2a and 6.5 (SD=2.9) in Experiment 2b. This suggests that when the task was clarified and especially when summary information was provided, participants' interventions judgments were closer to those arising from expected information maximisation and Bayesian inference.

There is no evidence that people were able to look more than one step ahead in this task though, with across-the-board worse fits for the *farsighted scholar, gambler* and *utilitarian* models and only 6/139 participants best fit by one of these models rather than the one-step ahead or null models. These were not the most successful participants, scoring an average of only 7.5, suggesting that the resemblance between their interventions and to those favoured by the two-step ahead expectancies was accidental.

In summary, comparing the "ideal" learner models shows that successful causal learners' actions and causal judgments are more closely related to the computational level problem of reducing uncertainty than those of maximising probability or utility.

Heuristic models

When comparing the full set of models, few of the participants were best described by either of the heuristic models. Nevertheless, these models fit relatively well despite their algorithmic simplicity, with BIC values in the range of the forgetful Bayesian models. Ignoring the expectancy based models, we can, similarly as previously for the "ideal" models, compare the heuristic models against the relevant null models (Table 3.3, last column). From this we can see that the better fitting heuristic overall is the *disambiguator*. However, more individual participants are better described as simple endorsers (61) than disambiguators (49), with the remainder being described by the *conservative random* null model. The majority (18/28) of those better described as

conservative random are in Experiment 1 and had average scores of only 6.4. Over Experiment 1, 2a and 2b, those described as disambiguators do slightly better than those described as simple endorsers $t(101.6) = -2.0, p = 0.04$. Disambiguators used complex interventions on 8.8% of trials (14.8% for Chain and fully-connected models) while simple endorsers rarely or never used complex interventions (1.3% of the time; 2% of the time for the Chain and Fully-connected models).

3.6 Discussion

Overall, our analyses suggest that the majority of people are highly capable active causal learners, both in terms of selecting useful interventions and in terms of learning from them. Having identified task confusion as the cause for many of the poorer performances in Experiment 1, we found that with an in-task reminder of the instructions in Experiment 2, almost all participants performed very well. Allowing participants to see the results of their past tests did not make a significant difference to performance but did significantly reduce task completion time. Since performance was already near ceiling in Experiment 2, we can see the quicker completion times suggesting that the history of past trials did make the task somewhat less demanding.

Simulations of efficient utility-, probability- and information-maximising active learning showed that starting with simple interventions and gradually switching to more focused "controlled" interventions made for an efficient interventional strategy. We see participants exhibiting this same pattern, starting with almost exclusively "simple" interventions and gradually using more "controlled" ones as they narrow down the space of possible structures. Participants' interventions were also somewhat sensitive to the structure being learned, with more "controlled" interventions being selected on the Chain and Fully-connected structures, where it was very hard to identify the correct causal structure without at least one controlled intervention. While participants were generally less inclined to select "controlled" interventions than the benchmark models, this is consistent with their learning being slower and more imperfect, as is reflected by our fitted models.

We can think of "simple" interventions as open-ended tests. They do not test any one hypothesis in particular and have multiple possible outcomes, each of which can be consistent with several different causal interpretations (e.g., if there are two activations following a simple intervention, these could result from a Chain, Collider, or

Fully-connected structure). However, simple interventions are powerful at first because they quickly reduce the space of likely models. In contrast, controlled interventions can be seen as more focused tests. They have only two possible outcomes and lend themselves to distinguishing unambiguously between two or three causal structures that perhaps differ by only one causal connection. This progression from open-ended to more focused testing gels with a picture of people as natural scientists, first exploring the space and identifying a candidate causal model, then progressively refining this with focused experiments. We found that the propensity to select controlled tests was closely linked to high performance, suggesting that only more sophisticated causal learners would progress from the exploratory stage to the stage of performing specific controlled hypothesis tests. The idea that controlled interventions are more cognitively demanding than simple ones is supported by research on complex control (e.g. Osman, 2011), where ability to recognize that one must simultaneously manipulate two variables to control a system is difficult for many people.

We found that participants had a strong tendency to mistake Chains for Fully-connected structures across both experiments. One reductive explanation for this is that some participants may have misunderstood the task demands, interpreting links as meaning that the parent node is a *direct or indirect* cause of the child node. However, the instructions were clear on this point, demonstrating the way in which fixing an in-between node "off" would block activation passing along a Chain. Instead, we conclude that this mistake is a marker for many participants' heuristic causal learning strategies. This is confirmed by the large number of participants whose actions and judgments are better described by the *simple endorsement* heuristic that systematically overconnects Chains rather than the *disambiguation* heuristic.

We compared computational models of efficient causal learning, driven by three plausible measures of intervention values: expected *utility gain*, *probability gain*, and *information gain*. Overall, the models driven by *information gain* (*scholars*) better fit the large majority of participants' interventions than models driven by *probability gain* (*gamblers*) or *utility gain* (*utilitarians*). This was particularly clear looking at the "ideal" models (without forgetting or conservatism). This means that, however participants were choosing their interventions and updating their beliefs, they were managing to do so in a way which broadly approximated the solution to the computational level problem of maximising information rather than that of minimising error or maximising utility.

Venturing one rung down the ladder from the computational level toward psychological process (Jones & Love, 2011; Marr, 1982), we explored "bounded" versions of our models.

We included forgetting and conservatism parameters capturing the idea that people might be biased in their learning by plausible memory and processing constraints. The fit of our models was greatly improved by inclusion of these parameters and including both parameters led to much better overall fits than including only one. The two parameters were correlated, supporting the idea that they complemented one another: e.g., the more forgetful a learner is about past evidence, the more conservative they need to be in their beliefs in order to be an effective learner. Therefore, these models provide an account of how moderate forgetting of old evidence paired with appropriate conservatism about existing causal beliefs can lead to effective active causal learning.

Allowing participants to draw and update models as they went along may have affected their learning, perhaps distracting them, or leading them to place more emphasis on earlier marked links. Furthermore, while we accept that the beliefs reported by participants at each time point are at best noisy markers of their actual beliefs about the true structure, we feel that these are largely unavoidable aspects of tracing beliefs throughout learning. We tried to minimize the extent to which eliciting beliefs distracted participants by making the step optional and hoping that participants would voluntarily record their beliefs as an aid to memory. It seems this was what most participants did, as links were drawn on 91% of tests, and neither varied wildly nor remained static from trial to trial. As a result we have been able to explore patterns of sequential causal learning in an unprecedented level of detail.

Taking another step toward the process level, we also looked at whether participants' actions could be reasonably captured by simple heuristics. We noted that *simple endorsement* (Fernbach & Sloman, 2009), based on local computation, could capture much of the behaviour of many participants. This may explain why so many participants judged the Fully-connected structure when the true structure was the Chain. However, some participants also performed the crucial controlling *disambiguation* steps which cannot be easily captured in a local computation framework. We operationalised this here as an alternative step occasionally performed at random. However, we note that a *disambiguator* type model has the potential to be refined by incorporating sequential dependence. For instance, a natural hypothesis is that disambiguation steps are most likely to be performed following ambiguous evidence (i.e. multiple activations). Another possibility is that learners are likely to perform disambiguation steps with the same node fixed on as they had fixed on for the step that generated the ambiguity. However, further refining the heuristic models in the current context is likely to make them increasingly indistinguishable from our expectancy based models. To confidently identify people's

heuristic strategies we will need to look at learning problems with a larger number of variables and the potential for larger divergences between heuristics and computational level models.

With these experiments and analyses we have begun the process of studying active causal learning behaviour, starting with a "simple" open ended experiment (Experiment 1), and more "controlled" follow up (Experiment 2). Having motivated and constructed models of participants' actions and judgments at computational and process levels, the next steps will be to come up with controlled tests that allow us to rigorously test some of these predictions. For example, an avenue of future work will be to look at the range of environments within which heuristic strategies are effective. We hypothesize that the extent to which one must disambiguate (or control for other variables) depends on how noisy, complex or densely causally connected the environment is. For more than around 5 or 6 variables, explicit calculation of expectancies becomes intractable while the calculations required by the active causal learning heuristics remain computationally trivial. In everyday life people have to deal with causal systems with many variables, far more than would plausibly allow explicit expectancies to be computed. One way people might achieve this is by performing an appropriate mixture of these "connecting" simple endorsement and "pruning" disambiguation steps.

3.7 Conclusions

In this chapter, we asked how people learn about causal structure through sequences of interventions. We found that many participants were highly effective active causal learners, able to select informative interventions from a large range of options and use these to improve their causal models incrementally over a sequence of tests. Successful learners were able to make effective use of "controlling" double interventions as well as "simple" single interventions, doing so increasingly as they narrowed down the hypothesis space. The large majority of participants behaved in line with our *scholar* model, choosing interventions likely to reduce their overall uncertainty about the true causal structure, rather than to increase their expected utility or probability of being correct. We found no evidence that people were able to plan ahead when choosing interventions. We also formulated bounded models allowing for forgetting and conservatism. These reveal that people exhibit recency when integrating evidence, but suggest that they may mitigate in part by conservative updating, preferring causal structures consistent or similar to their

previously stated beliefs. Finally, we identified simple endorsement and disambiguation as candidate heuristics for active causal learning.

Chapter 4

Children's active causal learning

"*Play is the answer to how anything new comes about.*"

— JEAN PIAGET

At the start of this thesis, I proposed that active causal learning is at the heart of cognitive development, implying that the rich causal theories we have as adults have their origins in discoveries we make through interventions as children. On this perspective, children must be able to intervene effectively and to interpret the causal implications of their actions. The modelling in Chapter 3 suggested that coming from a position of high uncertainty, it is valuable to perform simple interventions that perturb the world and make things happen, while more precisely targeted controlled interventions became more important as one narrows in on the true connections. Thus, we might expect a similar trajectory over development, with younger children focused more on making things happen, and older children gradually becoming more focused hypothesis testers. This chapter explores the development of intervention skills and structure judgments in five- to eight-year-olds, asking how choices and learning change and develop with age.

While there is a large developmental literature on self-directed learning, scientific and casual thinking, much of it is not grounded in the rational analysis nor tied to the CBN framework. Thus, this chapter begins by reviewing some of this developmental literature, linking it to my broader project. It then reports on an experiment in which five- to eight-year-olds learn through interventions in a simplified version of the adult tasks in Chapter 3, modelling intervention choices and structure judgments. Accordingly, this chapter represents the first formal model of children's intervention selection in learning structure relating multiple variables.

Developmental studies of causal learning (e.g. Bullock, Gelman, & Baillargeon, 1982; Gopnik, Sobel, Schulz, & Glymour, 2001; Shultz, 1982) have typically focused on determining the conditions under which children judge an event as causally efficacious. However, in learning about the world, what is at issue is often not just whether two specific variables are related, but the structure of the causal relations between a broader set of variables.

A number of studies have tasked adults with inferring the relations between sets of variables within the CBN framework (Fernbach & Sloman, 2009; Kushnir et al., 2010; Lagnado & Sloman, 2004, 2006; Sobel & Kushnir, 2006; Steyvers et al., 2003) with a smaller number having looked at intervention selection and learning from interventions (e.g. Bramley, Lagnado, & Speekenbrink, 2015; Lagnado & Sloman, 2004, 2006; Sobel & Kushnir, 2006; Steyvers et al., 2003). The CBN approach has also been adopted by developmental psychologists interested in explaining children's learning about causation (Gopnik, 2012; Gopnik et al., 2004). The majority of studies in this tradition have involved children learning whether an object possesses a particular causal power, usually on the basis of observing the experimenter's actions, such as whether an object makes a box light up and play a tune (e.g. Gopnik & Sobel, 2000; Schulz, Kushnir, & Gopnik, 2007; Sobel, Tenenbaum, & Gopnik, 2004). Relatively few studies have used tasks in which children themselves decide which interventions to carry out in order to discover the causal structure of a system (e.g., whether three variables are related in a Chain or a Fork structure). Such studies are particularly important because they can be used to assess young children's effectiveness in generating and testing hypotheses about the causal relations between sets of variables. Moreover, a key advantage of the CBN approach over other accounts of causal learning is that it captures this more complex type of learning, distinguishing between different causal paths as well as identifying variables' ultimate effects.

One study that did examine children's ability to learn causal structure through intervention is Schulz, Gopnik, and Glymour (2007, Exp 3), in which 4-to-5-year-olds intervened on a causal system involving a box with two gears. Children had to decide whether each gear moved by itself or was caused by the other. Children could remove each gear in turn from the box to test whether the other gear worked on its own when the box was switched on. They gave their answers about the relations between the gears by selecting from a set of anthropomorphised pictures of the gears depicting their possible relationships. Performance on this task was mixed. Children did not all reliably generate the right interventions to distinguish between the different possible relations that

might hold between the gears. Even those who did make appropriate interventions were not necessarily successful at identifying instances in which one of the gears caused the other one to move.

This study provides limited support for the claim that children can generate informative interventions and use the resultant information to distinguish between different causal structures. Not only was performance relatively weak, but children were only required to make judgments about the dependencies between pairs of variables, rather than to distinguish between multi-variable causal structures. Children gave their responses by pointing to pictures of the two gears that showed whether they turned themselves or one turned the other. There was a third variable that was important in the system (a switch) but children did not need to represent how its relation to the gears varied for the different causal systems and it did not feature in the pictures depicting causal relations between the gears. Thus, this study does not allow us to draw firm conclusions about whether children can use interventions to distinguish between, for example, Fork and Chain structures.

However, the findings of some other studies suggest that we should expect even very young children to be good at choosing appropriate interventions and using them to learn about causal systems (Bonawitz, van Schijndel, Friel, & Schulz, 2012; Cook, Goodman, & Schulz, 2011; Schulz & Bonawitz, 2007). Indeed, Schulz (2012) argues that the ability to select appropriate interventions and use the evidence generated from such interventions may be developmentally basic. In various studies, she has shown that young children will appropriately explore a causal scenario when given ambiguous information (Cook et al., 2011; Schulz & Bonawitz, 2007). In these scenarios, children's behaviour did suggest that they were trying to establish whether an object possessed a certain causal property. However, again children did not have to make interventions to disambiguate the structure of the relations between different variables, and then use this information to decide, for example, whether a system was a Fork or Chain.

A further study by Sobel and Sommerville (2010) tried to address this specific issue. Children viewed a box with four coloured lights, A, B, C, and D, and were told that some of the lights could make other lights turn on. The box was configured so that the relations between the lights took the form of either a Fork ($B \leftarrow A \rightarrow C, D$) or Chain ($A \rightarrow B \rightarrow C, D$) structure. Children could interact freely with the box by switching on lights and observing their effects. They were then asked a series of questions about the relations between pairs of lights. Sobel and Sommerville found that children performed above chance on these questions, which could be interpreted as indicating they were able

to use the information generated from their interventions to decide on the structure of the causal relations. There are two difficulties, however, with this interpretation. First, before children answered the questions, the experimenter pressed each of the buttons in turn and narrated what it did; arguably, this provided children with the answers to the test questions (indeed, children performed above chance, although less accurately, when just given this narration). Second, it was not clear that to answer correctly children needed to have an integrated representation of how the three variables in the system were related to each other, rather than just knowledge of pairwise relations. Indeed, Sobel and Sommerville do not analyse the answers that children give to the question of whether A makes C go in the case of the Chain, arguing that answers to this question are hard to interpret. However, by questioning children only about the other pairwise relations between A and B and B and C it is impossible to know whether children actually understood the nature of the overall causal structure.

The general point here is that we can distinguish between learning structurally local pairwise links and integrating such links to form a representation of causal structure. This distinction is important, because as we have seen, the complexity of causal learning scales rapidly as the number of variables increases. Learning localized pairwise relations is much easier than learning global structure (Fernbach & Sloman, 2009). However, if the other variables are ignored, this can lead to systematic mistakes about the global model. For example, when one connection "explains away" the dependence between two others, a local learning strategy, like the *Simple endorser* seen in the previous chapter, or Fernbach and Sloman's (2009) local computations approach, still attribute a connection between these two variables while a global strategy does not. A number of the studies of children's causal learning can be interpreted as studying children's learning of pairwise relations rather than global causal structure (Schulz, Goodman, Tenenbaum, & Jenkins, 2008; Sobel & Sommerville, 2009), meaning that we still have limited evidence about children's ability to learn causal structure.

Uncertainty as to whether children are adept at appropriately generating interventions and using them to learn causal structure comes from two sources. First, research on children's scientific learning has for many years suggested that younger children may have great difficulty generating appropriately informative interventions and learning the nature of relations between variables from the evidence generated by these interventions (e.g. Klahr & Dunbar, 1988; Klahr, Fay, & Dunbar, 1993; Kuhn, 1989; Schauble, 1996; Zimmerman, 2000). On the face of it, this body of findings seems at odds with recent findings from the CBN tradition. One possible explanation of the differing findings lies

in the role of the established causal theories (Griffiths & Tenenbaum, 2009), particularly in scientific learning studies (Tenenbaum et al., 2011). For example, pre-existing, and sometimes erroneous, beliefs can hamper children's ability to generate appropriate interventions and interpret statistical data (e.g. Amsel & Brock, 1996; Kuhn et al., 1988). Indeed, Schulz suggests that this type of factor, along with task complexity, may mask children's basic learning skills (Bonawitz et al., 2012; Cook et al., 2011), which may be better demonstrated in the tasks used in the CBN tradition in which domain-specific knowledge is of limited importance and the statistical evidence is simple.

4.1 Black box paradigm

However, the findings of a previous study by McCormack, Frosch, Patrick, and Lagnado (2015) provide a second reason for being unsure about children's ability to learn from interventions on a causal system. This study, like most of those in the CBN tradition, involved children learning about a novel mechanical system. The only relevant data for causal structure inference were statistical information provided through observing the experimenters' interventions on the system, and the temporal patterns of event occurrence. Children had to learn the causal structure of the system that was hidden inside a black box which had three separate shapes protruding from its top surface, that rotated (components corresponding to variables X_1, X_2, and X_3, see Figure 4.1). Across two experiments, children watched while an experimenter intervened on components in the system. In one experiment, the experimenter carried out interventions in which she disabled one of the shapes by preventing it from moving before moving each of the other shapes in turn. Children did not find it straightforward to use the patterns of evidence provided by these interventions to discriminate between causal structures, even when the system operated deterministically. Six-to-seven-year-olds were able to use the evidence from the more complex interventions to accurately infer when the system was one of the Chains. However, children younger than this could not do so, and even 7-to-8-year-olds were unable to use information from these interventions to accurately judge when the system was a Fork. McCormack et al. argue that children's difficulties may stem from integrating pieces of evidence provided across a number of separate observations of the causal system.

At first sight, McCormack et al.'s findings seem to be more consistent with the conclusions stemming from research on children's scientific learning that emphasized its limitations. However, this study did not provide children with an optimal opportunity

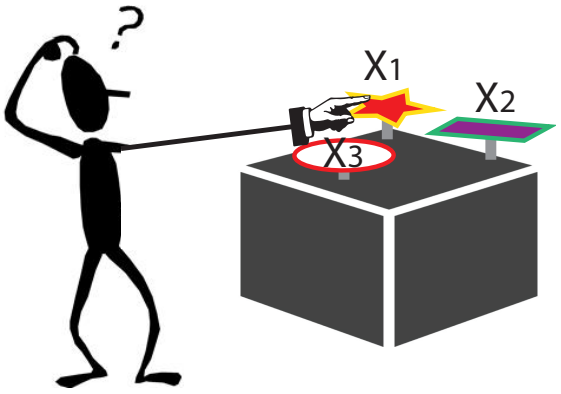


FIGURE 4.1: Schematic of the black box paradigm used for developmental experiment.

to demonstrate their abilities. Children watched while the experimenter made a series of interventions, rather than making the interventions themselves. Sobel and Kushnir (2006) argued that participants find it easier to learn causal structure when they decide what interventions to conduct, largely because this provides an opportunity for them to engage in more active hypothesis-testing (although see Lagnado & Sloman, 2004). A number of proposals in the literature (e.g. Markant & Gureckis, 2014; Markant et al., 2015) suggest an algorithmic basis for this effect — learners' choices are relevant to the hypotheses they are considering at the time of testing. If someone else chooses the interventions, there is no guarantee that they will be pertinent to what the participant is wondering about at the time. Moreover, children might be particularly likely to benefit from being allowed to explore how a system operates, in that hands-on interventions may ensure they stay engaged with the task.

In this study, we used a task very similar to that of McCormack et al. (2015), in which children had to decide whether a three-element causal system was a Fork ($X_2 \leftarrow X_1 \rightarrow X_3$), a 1-2-3 Chain ($X_1 \rightarrow X_2 \rightarrow X_3$) or a 1-3-2 Chain ($X_1 \rightarrow X_3 \rightarrow X_2$). Children intervened themselves in order to learn the box's hidden structure. Shapes on top of a box rotated either when children moved them by hand or they could be moved by rotating another shape that was causally connected to them (e.g., for the 1-2-3 Chain, spinning X_1 initiated the rotation of both X_2 and X_3, and spinning X_2 rotated X_3; all the shapes always moved simultaneously in the tasks to minimize temporal cues). Children had to select and carry out a series of interventions; these could be simple interventions in which they made one of the three shapes spin, or they could be more complex interventions in which children prevented one of the three shapes from moving by disabling it, and then spun one of the other two shapes. Note that we were not attempting to faithfully recreate a free-play situation because it was important for our analyses that we were

able to exhaustively categorize children's actions on the system. Although they were completely free to choose their interventions, the only actions children could carry out were interventions on the system. Furthermore, it was made clear to children that their job was to learn the causal structure of the system, and that they could not make an unlimited number of interventions. This allowed us to look in our modelling work at the efficiency with which children produced informative interventions.

We examined two aspects of performance: the nature of the interventions that children selected and their causal structure choices. Not all interventions provided useful information to discriminate between the three possible causal structures, which allowed us to examine whether the tendency to choose informative interventions changes with age. We also examined whether there was any relation between the quality of children's interventions and the likelihood that children chose the correct causal structure at test. It is possible to try to examine these issues without formal modelling (see Sobel & Kushnir, 2006), by, for example, simply distinguishing between two broad classes of informative and non-informative interventions. However, we chose to model children's learning in a Bayesian framework. Doing so has two key advantages: First, it allows us to properly assess whether there are developmental changes in the extent to which children resemble an idealized Bayesian learner. This is important because, within the currently dominant CBN tradition, young children's learning is often characterized as approximating such an ideal, particularly with regard to causal learning from statistical information (e.g. Gopnik, 2012; Gopnik & Wellman, 2012). Formal modelling allows us to assess the extent to which this characterization is appropriate by assessing children's performance against the standards set by the Bayesian tradition itself. Second, although in this study we can (and do) classify interventions broadly as informative or non-informative, the learning task itself is sequential. This means that how informative an intervention is depends on what children have already observed, and what they can remember about such observations. However, figuring out the informativeness of each intervention that a participant makes on a trial-by-trial basis would be a formidable task without a formal model. Indeed, without such a model it is hard to see how one would operationalise the notion of informativeness under such circumstances. Our Bayesian model allowed us to capture the sequential nature of the learning task, by assuming that the most informative interventions were those that maximally reduced uncertainty about which was the correct hypothesis at any particular point in the learning sequence, given some level of forgetting.

4.2 Experiment 3: Children's interventional learning

4.2.1 Method

Participants

Seventy-seven children participated, from three different school years: 21 5-to-6-year-olds (M = 72 months, Range = 64-80 months), 31 6-to-7-year-olds (M = 86 months, Range = 80-93 months) and 25 7-to-8-year-olds (M = 98 months, Range = 93-103 months). Children were tested individually in their schools.

Materials

The study used a wooden box, 41 cm (long) x 32 cm (wide) x 20 cm (high), which had an on/off switch at the front. There were three different coloured lids for the box. Two of these had three coloured/patterned shapes (e.g., circle, rectangle, star) inserted on its surface that rotated independently on the horizontal plane; a separate lid was used in pretraining and had only two shapes (see Figure 4.2).

The colour and shapes of the components were varied across participants and causal structures. On each of the two lids used in testing, the three shapes formed an equilateral triangle of sides 24 cm. Each shape had a small hole that aligned with a hole in the lid of the box. There was a miniature red-and-white "Stop" sign affixed to a metal rod that could be inserted through the hole on any shape into the corresponding hole in the box, preventing it from moving. Each of the shapes could be rotated by hand; the rotation of the other shapes was controlled by a laptop hidden inside the apparatus that participants were unaware of. A set of photographs was used during the learning phase that participants used to indicate which intervention they were going to make; these photographs depicted each shape on the box and in addition there were photographs of each of the shapes alongside the stop sign. Photographs of the whole box with its shapes depicting three possible causal structures were used at test for children to indicate their judgment of the causal structure: one Fork and two Chains (i.e. depicting $X_2 \leftarrow X_1 \rightarrow X_3$, $X_1 \rightarrow X_2 \rightarrow X_3$, or $X_1 \rightarrow X_3 \rightarrow X_2$). The photographs for use at test were overlaid with pictures of hands to indicate causal links (following Frosch et al., 2012).

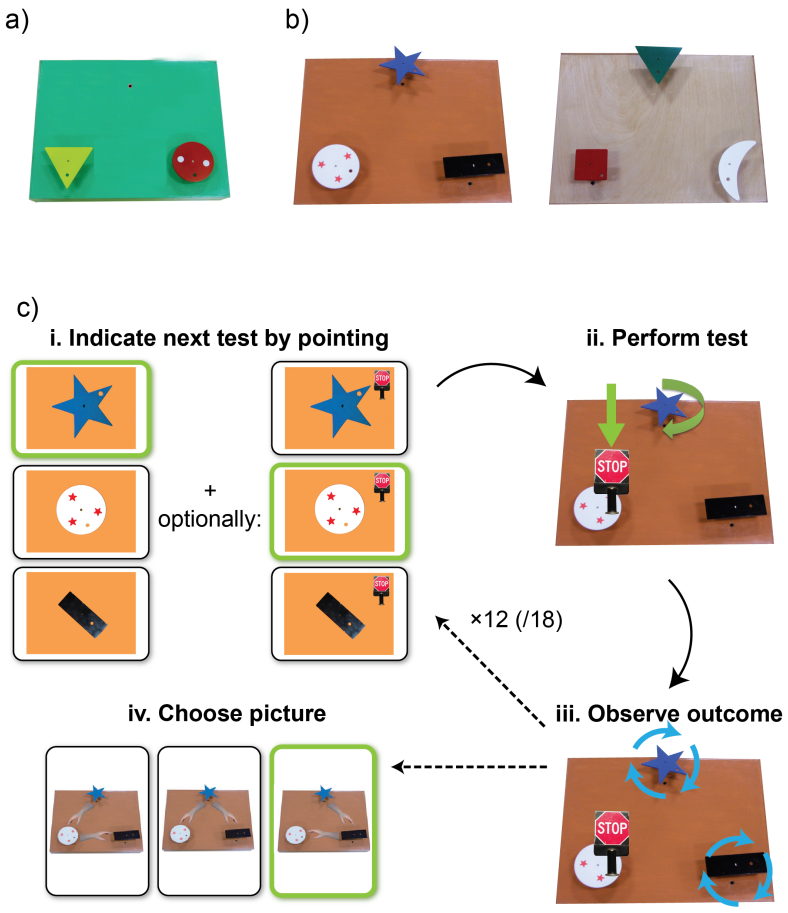


FIGURE 4.2: a) The box lid used in training; b) The two box lids used during testing (counterbalanced between common-cause and Chain trials). Procedure: c) i. Participants indicate (optionally) which shape to block with the stop sign and which shape to spin; ii. Perform the action(s) they chose; iii. Observe which shapes spin as a result of their test; iv. After 12 (or 18) tests, they point to the card showing how they think the machine works. Green arrows and highlighting show participants' actions on an example trial.

Procedure

Children completed two test trials, one Fork and one Chain (order counterbalanced). There was a pretraining phase that ensured children knew what their task was and how to give their answer at test. The pretraining procedure used a lid on the box that had only two coloured shapes inserted on its surface; its purpose was to demonstrate that some shapes caused others to move but that the stop sign could be used to prevent a shape from moving. Children were initially asked to name the colours of the shapes to ensure that they would know to which shapes the experimenter was referring, and the experimenter drew children's attention to the on/off switch at the front, set at the "off" position. She then switched the box on and manually rotated one of the two shapes

(X_1). This had no effect on the other shape (X_2), which remained stationary, and the experimenter pointed this out to children. She then rotated the other shape (X_2), which resulted in the first shape (X_1) simultaneously rotating. She explained to children that "Some shapes are made to move by others". The experimenter then switched the box off and introduced children to the stop sign, which she inserted into X_1 to stop it from moving, saying "See this stop sign, it can be used to stop a shape from moving, see the [colour X_1] one cannot move now". She then switched the box on again and rotated X_2, which this time had no impact on the movement of X_1 because it was prevented from moving by the stop sign. Following this, the lid was removed from the box, and replaced by a different coloured lid with three different shapes for the first test trial.

Children were asked to name the colours of the three shapes and were told that their job was to figure out how the box worked. They were introduced to the three test pictures depicting the three different causal structures with the experimenter saying: "In a moment I will ask you to figure out how the box works, but first I want to show you some pictures of the box which show different ways in which the box may be working. Only one of them is right and you've got to work out which is the right one. It won't change half way through, and it is definitely only one of the pictures. You'll have to use your detective skills to work out which picture shows what the box does." The experimenter described each of the three pictures (e.g., "In this picture, the red one makes the blue one go, and the blue one makes the white one go, and the hands show that"). Following these three descriptions, children were then asked a set of three comprehension questions. For each Chain picture, the experimenter asked "Can you show me the picture where the [colour X_1] one makes the [colour $X_{2/3}$] one go and the [colour $X_{2/3}$] one makes the [colour $X_{3/2}$] one go?", and for the Fork picture "Can you show me the picture where the [colour X_1] one makes both the [colour X_2] one and the [colour X_3] one go?). The majority of children answered these questions correctly first time, but if they did not answer all three questions correctly, the experimenter repeated the initial descriptions and asked the comprehension questions again. This procedure was repeated again if necessary.

Following this pre-training, the experimenter said: "I am going to switch the box on now and I want you to figure out how the box works." Children were told that they could do one of two things (order counterbalanced): either "You can move a shape to see if it makes other shapes move" or "You can stop a shape from moving by putting the stop sign in and then see what happens when you move another shape". It was explained to children that, before they carried out each intervention, they had to point to a card

indicating what they intended to do. The experimenter said: "Before you try anything on the box I want you to point to one of these cards. This card means you want to spin the [colour] one, and you point to this card if you want to stop the [colour] one. See we also have the cards for spinning and stopping the [colour and colour] ones. So each time you want to do something, you point to one of these cards first." Children were told that they had 12 goes "to start with" and that each time they moved a shape counted as one go. It was made clear that using the stop sign did not count as a go by itself; children had to then in addition move one of the other shapes. The procedure with cards was used to ensure children interacted with the box in a controlled way and to make clear that they could not make an unlimited number of interventions. It also ensured that all children made a fixed minimum number of interactions before attempting to answer the test question. Children were told that they did not need to keep track of the number of goes that they had with the box, as the experimenter would count this for them.

Before children began, the experimenter said "Remember, you've got to figure out which picture shows how this box really works." She then demonstrated what happened when shape X_1 was moved, which was that the other two shapes also moved simultaneously, and pointed out that they didn't know yet "which ones make other ones go". Participants were subsequently allowed to make interventions on the box by first selecting the appropriate card and then making the intervention. So, for example, if they wanted to see what happened when X_3 was moved if X_2 was disabled, they had to point to the card depicting X_2 with the stop sign in it, and then to the card depicting X_3. They then carried out their intervention.

After the participants had completed 12 interventions, the experimenter said "You have had your 12 goes now — do you want to choose which picture you think shows what the box did, or do you want to have another 6 goes?" The majority of participants opted to choose after 12 interventions. Children completed a short filler task (a paper-and-pencil maze) in between the Fork/Chain trials. It was made clear that the second box might work in the same way as the first box or it might work in a different way. The second box always had a lid of a different colour and different shapes.

4.2.2 Results

In both trials, 69/77 participants stopped after 12 interventions. The remaining 8 opted for an additional six interventions in one or other trial. Of these, 4 participants opted for the additional six interventions on both trial types. Initial data analyses examined

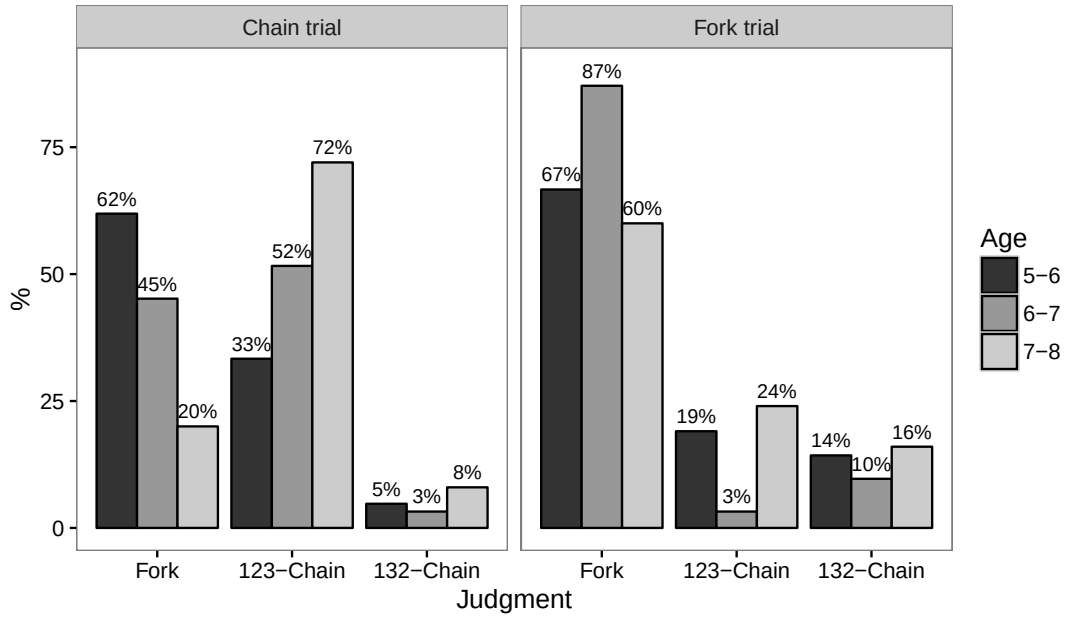


FIGURE 4.3: Percentage of responses for each causal structure as a function of age group and trial type. Correct responses to the Chain are denoted as *123-Chain*

participants' responses for each of the two trial types. Figure 4.3 shows the percentage of participants who chose each response type for each trial type. The majority of participants in each group, except for the youngest group, chose the correct answer for the Chain trial. The majority of participants in all groups chose the Fork response for the Fork trial. χ^2 tests showed that each group of participants chose the correct response more often than chance, all $ps < .01$, except for the 5-to-6-year-olds, who did not select the Chain more often than chance. This group tended to select the Fork response for both structures. Performance on the Chain structure was associated with age, $\chi^2(2) = 6.91, p < .05$, with the number of correct responses improving with age. Performance on the Fork structure was marginally significantly associated with age, $\chi^2(2) = 5.66, p = .056$, although in this case the 6-to-7-year-olds gave more correct responses than each of the other groups.

Analysis of interventions

Subsequent analyses examined the nature of participants' interventions on the system. We initially discriminated between whether an intervention was informative or not, given the three possible causal structures. There were three interventions that were never informative: $\text{Do}[X_1 = 1], \text{Do}[X_2 = 1, X_3 = 0]$, and $\text{Do}[X_2 = 0, X_3 = 1]$. Potentially informative interventions were $\text{Do}[X_1 = 1, X_2 = 0]$, $\text{Do}[X_1 = 1, X_3 = 0]$, $\text{Do}[X_2 = 1]$, $\text{Do}[X_3 = 1]$, $\text{Do}[X_1 = 0, X_2 = 1]$ and $\text{Do}[X_1 = 0, X_3 = 1]$. We also classified interventions

as simple or complex: Do[$X_1 = 1$], Do[$X_2 = 1$] and Do[$X_3 = 1$] were classified as simple and those involving initially disabling one of the components before moving another component as complex. Table 1 shows the percentage of times that participants in each age group chose each of these interventions. The most popular intervention tended to be to be Do[$X_1 = 1$], which, although it was uninformative, did make all of the three shapes spin. Propensity to select a complex intervention increased significantly with age $F(2, 74) = 7.22, p < .002, \eta^2 = 0.16$, with 7-to-8-year-olds the most likely to pick the complex interventions (65% of the time, compared to 46% for 5-to-6-year-olds and 45% for 6-to-7-year-olds).

TABLE 4.1: Percentage of Time Participants Chose Each Intervention

	Informative						Uninformative		
	$X_2=1$	$X_1=0$ $X_2=1$	$X_3=1$	$X_1=0$ $X_3=1$	$X_1=1$ $X_3=0$	$X_1=1$ $X_2=0$	$X_1=1$	$X_2=0$ $X_3=1$	$X_2=1$ $X_3=0$
5-6 years	17.9	9.5	12.9	5.1	9.6	8.3	23.3	8.2	5.4
6-7 years	17.5	9.2	14.2	8.2	8.2	9	21.8	6.4	5.6
7-8 years	10.8	11.3	9.4	11.1	14.8	13.5	14.2	6.9	8

Note: Collapsed across common cause and causal chain trials.

We examined whether participants chose informative interventions more often than chance by conducting a one-sample t-test with a test value of 0.67, given that $\frac{2}{3}$ of the 9 possible interventions were informative. Only the 7-to-8-year-olds were significantly more likely than chance to select informative interventions, $t(24) = 2.83, p < .01$, both $ps > .10$ for the younger groups. A logistic regression showed that the proportion of informative interventions significantly predicted the probability of a participant getting the Chain trial correct $z = 2.73, p < .01$ (see Table 4.2), but this was not the case for the Fork trial $z = -0.62, p > .5$. One potential explanation for the latter finding is that the children were overall more likely to select the Fork, doing so 56% of the time. Thus, some of the correct responses on the Fork test trial are likely to have been made by the weaker participants purely in virtue of their favouring the Fork structure.

4.2.3 Modelling Interventions

So far, we have looked at proportion of informative intervention choices without considering the sequential nature of the task or whether and how efficiently children produced a set of informative interventions sufficient to discriminate between causal structures. A child who did not produce such a set but repeatedly produced a single informative intervention would score 100% on this measure. Moreover, how useful an intervention is

depends on what the learner already knows (in this case what they have already learned from their previous interventions). For example, $\text{Do}[X_1 = 1, X_2 = 0]$ and $\text{Do}[X_3 = 1]$ are both informative interventions in this task provided you do not know anything yet. But suppose you have already performed $\text{Do}[X_1 = 1, X_3 = 0]$ and observed that this made X_2 spin. This evidence effectively rules out the 1-3-2 Chain leaving only the 1-2-3 Chain and the Fork as possibilities. Now, on subsequent trials, performing $X_3 = 1$, or repeating $\text{Do}[X_1 = 1, X_3 = 0]$ will not tell you anything new, as both of these interventions simply distinguish the 1-3-2 Chain from the other two. To capture how efficiently children's intervention choices allow them to home in on the true structure we can analyse the interventions sequentially by looking at how effectively these interventions reduce uncertainty, assuming initially children are perfectly able to remember past outcomes and integrate new information.

TABLE 4.2: Experiment 3: Regression analyses

Dependent	Parameter	Estimate	SE	Odds ratio	Z	$P(>z)$
$\text{P}_{\text{correct}}$(Chain)	Intercept	-4.77	1.8		-2.65	.008**
	% Informative	7.37	2.7	1587	2.73	.006**
$\text{P}_{\text{correct}}$(Fork)	Intercept	1.87	1.48		1.27	.21
	% Informative	-1.32	2.15	0.267	-0.62	.54
$\text{P}_{\text{correct}}$(Chain)	Intercept	-0.81	0.75		-1.09	.278
	Efficiency	1.22	0.92	3.39	1.32	0.186
$\text{P}_{\text{correct}}$(Fork)	Intercept	3.39	0.99		3.39	< .001***
	Efficiency	-3.31	1.27	0.037	-2.62	0.009**
$\text{P}_{\text{correct}}$(Chain)	Intercept	-4.43	1.74		2.53	.012*
	Quality	7.98	3.06	2921	2.61	.009**
$\text{P}_{\text{correct}}$(Fork)	Intercept	1.46	1.4		1.04	.3
	Quality	-0.78	2.24	0.46	-0.35	.73

Note: * p <.05, ** p <.01, *** p <.001. Three separate analyses of predictors of performance on the causal chain and common cause trials: *Percentage Informative Interventions, Efficiency, and Intervention Quality.*

To do this, we defined a participant's subjective uncertainty about the true structure at a given time point as the information entropy $H(M)$ (Shannon, 1951) of their posterior distribution $P(M)$ over the three possible structures, given the data they had seen so far (Equation 3.9). Every time the child observed new evidence $\mathbf{d}^t$ associated with their chosen intervention $\mathbf{c}^t$, this distribution was updated using Bayes rule and the likelihoods $P(\mathbf{d}^t|M;\mathbf{c}^t)$ for observing that outcome out of the possible outcomes $\mathbf{d}^t \in \mathcal{D}_{\mathbf{c}}$ for each structure $m \in M$ giving posterior probabilities $P^t(M|\mathbf{d}^t;\mathbf{c}^t)$ (see Equation 3.3). Because the box worked in a deterministic way, the likelihoods were always 0 (if the outcome was impossible given that structure and intervention) or 1 (if that outcome was to be expected given that structure and intervention). If an outcome had zero likelihood

under one structure, then that structure's posterior probability would go to zero once a participant saw that outcome. By doing this we were able to compute the expected information gain $\text{Eig}^t(\mathbf{c}|m)$ for each intervention chosen by participants (see Equation 3.10), and rescale this by the maximum achievable expected increase in information over the different interventions at that time point. This gave a measure of the overall efficiency of the intervention choices made by each child for facilitating their identification of the true structure

$$\text{Efficiency}_{\mathbf{c}^t} = \frac{\text{Eig}^t(\mathbf{c}|m)}{\max_{\mathbf{c}' \in C} \text{Eig}^t(\mathbf{c}'|m)}. \tag{4.1}$$

Because of the deterministic nature of the task, in fact all of the children generated enough information with their interventions for their uncertainty to go to zero before the end of the trial so intervention efficiency was simply calculated for the interventions up until the point that their posterior uncertainty reached zero. Figure 4.4 shows an example of how the model worked using a real set of interventions; it also depicts how these interventions were categorized. We established chance level interventional efficiency by simulating the task 1000 times with randomly selected interventions, finding average chance efficiency levels of .48 for the Chain structure and .43 for the Fork.[1]

For all age groups, for both structures, children's interventions were significantly much more efficient than the chance level (mean efficiencies for Chain and Fork respectively: 5-to-6-year-olds = 0.71 and 0.66, 6-to-7-year-olds = 0.79 and 0.63, 7- to-8-year-olds = 0.82 and 0.80; all $ts > 6, ps < .001$). Children's efficiency for the Fork changed significantly with age, $F(2, 74) = 4.47, p =< .02, \eta^2 = 0.11$, but there was no effect of age on efficiency on the Chain trials, $F(2, 74) = 1.14, p = 0.32, \eta^2 = 0.03$. Unlike proportion informative interventions, efficiency did not predict accuracy on the Chain (see Table 4.2), and was in fact negatively related to accuracy on the Fork ($z = -2.62, p < .01$).

While proportion informative interventions did not take into account the sequential nature of the task, arguably interventional efficiency has the opposite shortcoming. By assuming, implausibly, that children have a perfect memory for the outcomes of their previous interventions and perfect ability to make inferences from this information, it ignores what they do on subsequent interventions once they have, in principle, enough

[1]This corresponds to getting enough information to identify the true structure after an average of 3 random interventions when the true structure is a chain and 4.5 random tests when the true structure is the common cause. The chain is somewhat quicker to be identifiable by chance because sometimes it can be identified from a single intervention (e.g. $\text{Do}[X_2 = 1]$ allows identification of the 1-2-3 Chain), while the common cause always requires a minimum of two interventions.

Test	Intervention	Result	Informative	Prior	Efficiency	Prior (25%)	Quality
1	X2=1, X1=0	X3	Y		1		1
2	X2=1, X1=0	X3	Y		0		.42
3	X1=1, X2=0	X2	Y		1		1
4	X2=1		Y		/		1
5	X3=1, X1=0		Y		/		1
6	X3=1, X1=0		Y		/		.66
7	X1=1, X2=0	X3	Y		/		1
8	X3=1		Y		/		1
9	X1=1, X3=0	X2	Y		/		1
10	X1=1	X2, X3	N		/		0
11	X1=1	X2, X3	N		/		0
12	X2=1, X3=0		N		/		0
				F 123 132		F 123 132	
	Proportion *complex*		**Proportion *informative***		**Mean *efficiency***		**Mean *quality***
	.66		**.75**		**.66**		**.64**

FIGURE 4.4: Interventions selected by a 6- to 7-year-old, in the Fork trial. From left to right, columns show: 1. Test order. 2. Selected intervention. 3. Which, if any, shapes spun as a result. 4. Whether the intervention was generally informative. 5. A learner's prior given perfect memory and integration of previous tests (Bars are (F)ork, (123) Chain and (132) Chain). 6. The corresponding efficiency of the intervention in allowing identification of the Fork. 7. A learner's prior given 25% forgetting. 8. The corresponding quality of each intervention.

information to potentially identify the correct structure. An inspection of the modelled data found that children obtained sufficient information for certainty after an average of only 2.75 interventions; this means that our measure of efficiency ignores a large proportion of the data and makes no allowances for noise, forgetting, or uncertainty in learning. A more balanced way to assess the quality of participants' interventions is achieved by adding some noise, encapsulating the idea that learning is likely to be somewhat leaky or error prone.

We augmented our Bayesian learning model so that, after each test, some proportion of what was learned previously was "forgotten".[2] This was achieved by mixing a uniform

[2]There are numerous ways to model forgetting (e.g. Wixted, 2004) but a reasonable high-level approach is to assume children forget random aspects of their priors, leading to a net "flattening" of their subjective priors going into each new intervention. We remain agnostic about whether this parameter captures cognitive forgetting, or more generalized sources of error and uncertainty in children's integration of information.

distribution in with the posteriors, with the proportion determined by a forgetting rate $\gamma \in [0, 1]$

$$P^t(M|\mathbf{c}, \mathbf{d}) \propto P(\mathbf{d}^t|\mathbf{c}^t, M) \left[(1-\gamma)P^{t-1}(M) + \frac{\gamma}{3}\right] \tag{4.2}$$

(similar to Equation 3.11), and using this as the prior for next intervention. This procedure was carried out for each of the learner's 12 (or 18) tests. This means that previously-ruled-out alternatives gradually regained some probability mass, while more likely options became a little less favoured. The quality of each intervention was then calculated based on the extent that it reduced uncertainty across these distributions, compared to an intervention that would have maximally reduced uncertainty. This method captures the idea that continually repeating a particular intervention is less useful than selecting a complementary mixture of different interventions while also allowing that real world learners are likely to forget, ignore, or make mistakes about the evidence they have seen previously, meaning that revisiting previous interventions is not useless.

The exact level of "forgetting" in the model turned out not to be particularly important. We found qualitatively the same results setting it to 10%, 25%, 50%, 75% and 90% although the results were clearer for the lower levels of forgetting. Here we report results assuming 25% "forgetting" after each test. We established chance levels of intervention quality, again through simulation over 1000 trials. The 5-to-6-year-olds' intervention quality was not significantly above chance on either the Chain or Fork, 6-to-7-year-olds were above chance on the Chain, $t(30) = 2.88, p < 0.01$, and marginal on the Fork $t(30) = 1.73, p = 0.09$, while the 7- to-8-year olds' intervention quality was above chance for both trials (Chain: $t(24) = 2.46, p < 0.02$, Fork: $t(24) = 4.87, p < .001$). Averaged over trial types, we found that intervention quality improved with age, $F(2, 74) = 4.03, p < .03, \eta^2 = .10$, with 7-to-8-year-olds significantly more efficient that 5-6-year-olds, $p < .01$, but no significant difference between 5-to-6-year-olds and 6-to-7-year-olds. Breaking this into responses for the two structures, regardless of forgetting rate, intervention quality was a significant predictor for correct identification of the Chain structure, $z = 2.61, p < .01$, but not for the Fork structure (see Table 4.2).

4.2.4 Discussion

Our findings provide important information about developmental changes in children's ability to learn causal structure through intervention. Children's ability to learn a

Chain structure improved with age, with the youngest children unable to identify this structure above chance. However, we need to consider why the 5-to-6-year-olds identified the Fork structure as accurately as the 7-to-8-year-olds. Our view is that the good performance on this second trial type is due to a tendency even amongst the youngest children to assume that, when events happen simultaneously, the underlying structure is a Fork. Previous studies have found that both children and adults make use of this simple temporal heuristic when they observe a three-variable system with this sort of temporal schedule (Burns & McCormack, 2009; Fernbach & Sloman, 2009; Lagnado & Sloman, 2006). Indeed, McCormack et al. (2015) demonstrated that children will use this type of temporal heuristic even when faced with contradictory statistical information provided either through observing the operation of a probabilistic causal system or through observing the effects of interventions on a deterministic system. Thus, the good performance of the younger children on the Fork structure is likely to reflect use of this temporal heuristic rather than use of statistical information derived from interventions on the system. This would also straightforwardly explain the lack of a relation between intervention quality and performance on the Fork structure.

The analyses of children's intervention choices provide insights into why performance improved developmentally on the Chain trial. Interventions could be initially classified as informative or non-informative, given the three possible causal structures. Over all trials, unlike the oldest group, younger groups of children did not choose informative interventions more often than chance. It proved fruitful, though, to further examine intervention choices and how these related to performance by modelling intervention selection. The initial analysis of how efficient participants were at producing a set of interventions that could, in principle, discriminate between the different hypotheses showed that all groups of children produced such a set more quickly than would be expected if they were simply choosing between interventions at random. This means that even the youngest children had the evidence available to them to make the appropriate causal inferences. However, intervention efficiency was not a predictor of performance. This demonstrates that selecting interventions that are, in fact, disambiguating, is not sufficient for good performance. Children may forget or fail to make use of what they have observed, and the subsequent interventions they make may also influence their judgments.

Our modelling work suggested this was indeed the case, because our measure of the quality of children's interventions that took into account the complete sequence of interventions predicted performance on the Chain structure, under the assumption that there

was some degree of forgetting. Moreover, unlike efficiency, intervention quality improved with age, with older children being more likely to consistently choose interventions that would help disambiguate the causal structures, given what they had already observed. These results indicate that with development, children become more discerning in their choice of interventions, and this has an impact on their causal structure learning.

How do our findings fit with what is already known about developmental changes in children's use of interventions to learn about causal systems? In designing our study, we sought to ensure that domain-specific knowledge was not relevant for task performance. However, this did not rule out children exploiting a type of pre-existing, albeit domain-general, heuristic about the nature of the causal system, namely that when multiple events occur immediately following an intervention the underlying structure is likely to be a Fork (McCormack et al., 2015). When the evidence generated from interventions was consistent with this assumption (i.e. in the Fork trial), even young children performed well. However, younger children had difficulty discarding this assumption on the basis of the contradictory evidence provided by their interventions. This is consistent with evidence from scientific learning literature that indicates that children have difficulty discarding a pre-existing hypothesis and may routinely ignore statistical evidence that fails to support such a hypothesis (Amsel & Brock, 1996; Kuhn et al., 1988).

Furthermore, an inspection of developmental changes in the pattern of children's intervention choices (Table 4.1) yields some further interesting additional parallels with findings from the scientific learning literature. Our task is very different to those used in research on children's scientific learning: it is simpler and the children that we tested are younger than those typically used in such studies (although see Koerber, Sodian, Thoermer, & Nett, 2005; Piekny & Maehler, 2013). Nevertheless, some of our findings confirm broad developmental patterns that are well-established in that research. Younger children tended to prefer making the $\mathrm{Do}[X_1 = 1]$ intervention, and did so repeatedly. This intervention is the most causally effective (it makes all the events happen), but does not discriminate between the three available hypotheses. However, it reinforces any existing hypothesis that the causal structure is a Fork by providing the temporal pattern of all events happening simultaneously. Young children's preference for this intervention has parallels with demonstrations in the scientific learning studies that show that children attend most to the variable already believed to be causal, focus more on producing an effect than on generating disambiguating evidence, and produce evidence that is consistent with their existing hypothesis rather than seeking to disconfirm it (Klahr & Dunbar, 1988; Klahr et al., 1993; Kuhn, 1989; Schauble, 1990, 1996).

Although younger children's patterns of interventions led to poorer performance, it is interesting to note that recent formal analyses have demonstrated that whether their type of approach should be viewed as inefficient depends on the learning context. First, the tendency to intervene on variables already believed to be causal in order to confirm an existing hypothesis is not necessarily always the wrong strategy. This type of strategy has been shown to be rational under the assumption that causal connections in the world are sparse (Navarro & Perfors, 2011), meaning that competing causal hypotheses do not generally share the same effect variables. In such circumstances "positive tests", operationalised as intervening on the variable thought to be the root cause (Coenen et al., 2015) are highly diagnostic. Hence, younger children's pattern of interventions could be interpreted as due to a tendency to act in ways that have proved an effective general-purpose strategy for learning causal relationships in the past, despite being inappropriate in the current learning context.

Second, we also found that younger children were less likely than older children to produce the more complex interventions that involved disabling one of the components in the system. This type of intervention can be particularly informative because it can be used to exclude a variable as being necessary for production of an effect. However, separate Bayesian modelling work with adults has demonstrated that producing simple rather than complex interventions is not always an inefficient strategy. Bramley, Lagnado, and Speekenbrink (2015) show that simple interventions tend to be more informative than complex interventions with respect to a broader hypothesis space (e.g. all possible 3-variable causal models), with more complex interventions becoming more useful once the space of possibilities narrows to favour a smaller number of hypotheses (e.g. those that differ by just a single edge). In our task, children had to discriminate between just three competing hypotheses, so it is one in which complex interventions are likely to be useful. In summary, the observed developmental changes can be interpreted as supporting the idea that while younger children used simple strategies that may be effective in other contexts, older children were better able to adjust their learning strategy in a way that was appropriate the task — i.e. to use a "control of variables" strategy (Chen & Klahr, 1999; Dean Jr & Kuhn, 2007) whereby confounding variables are experimentally controlled.

4.3 General Discussion

Experiment 3 examined children's causal structure learning under circumstances in which they selected and carried out interventions on a simple three-variable causal system. To the best of our knowledge, the analyses reported here of its data constitute the first attempt to model the quality of children's interventions when learning causal structure within a Bayesian framework. Our findings regarding children's interventional learning varied depending on whether children were learning a Chain or a Fork structure. With regard to the former, there were clear developmental improvements not only in terms of accuracy of structure learning, but in terms of the quality of the interventions that children produced as assessed by our modelling. The key advantage of the modelling is that it provided us with a quantitative measure of the *quality* of children's interventions, allowing us to examine the informational content of children's interactions with the devices. This Bayesian measure of intervention quality predicted performance. Put simply, the findings suggest that with development, children increasingly resemble an idealized Bayesian learner, although we note that the best predictor of performance from our modelling results was a measure of interventional quality that assumed substantial noise in the Bayesian learning process.

The same pattern of findings did not obtain for the Fork structure, and the most plausible interpretation of this is that younger children's inferences in this task were based on a simple temporal heuristic (e.g. "assume a Forking structure if multiple effects occur simultaneously") rather than on use of statistical information provided from interventions. Use of such temporal heuristics is widespread in both children's and adults' causal structure learning (Burns & McCormack, 2009; Fernbach & Sloman, 2009; Lagnado & Sloman, 2004, 2006; White, 2006b), with McCormack, Bramley, Frosch, Patrick, and Lagnado (2016) demonstrating that younger children's causal structure inferences are highly influenced by the temporal pattern of events. Their findings are consistent with those from the current study, insofar as those authors also found no developmental improvements in the likelihood that children would give a Fork judgment under circumstances in which all events happened simultaneously. Children's tendency to recruit temporal heuristics is likely to be due to the heuristics' low demands on information processing in comparison to using statistical information (Fernbach & Sloman, 2009). For example, in the current study, use of such a heuristic would have been based on the observation of a single intervention: the temporal pattern of events following $\mathrm{Do}[X_1 = 1]$.

This intervention was the most common one made by both the youngest groups; we interpreted this as suggesting that these children focus on producing an effect rather than systematically testing the competing hypotheses, and in doing so are provided with evidence (i.e. the temporal pattern of events) that they take to be consistent with their existing hypothesis. Younger children were also less likely to disable components in the system, suggesting that they were less likely to try to exclude any variables. Although younger children's interventions on the system had these characteristics, all children produced a set of interventions that could in principle have allowed them to correctly judge the causal structure. However, the Bayesian analysis proved useful in establishing that simply initially producing interventions that could potentially disambiguate the causal structure was not predictive of good performance. Rather, children's performance was related to how informative their interventions were as they moved through the task sequentially, with the Bayesian modelling capturing the idea of evolving beliefs guiding a sequence of intervention choices.

4.4 Conclusions

The findings of this study point to two clear directions for future work in this area. First, the fact that children become increasingly Bayes-efficient information-seekers in their causal learning raises the question of what cognitive changes underpin this developmental shift (see Lucas, Bridgers, et al., 2014, for recent work in this direction). While we did not attempt to model psychological processing explicitly here, models based on approximate Bayesian inference that attempt to be more psychologically plausible (Bramley, Dayan, Griffiths, & Lagnado, 2017; Bramley, Dayan, & Lagnado, 2015; Dasgupta et al., 2016; Kemp, Tenenbaum, Niyogi, & Griffiths, 2010; Sanborn et al., 2010) may play a role in addressing this question. Importantly, the developmental improvements found in Experiment 3 highlight the need for Bayesian models that do not just capture idealized learning but can accommodate, and potentially explain, developmental changes in the quality of children's causal learning. Explaining developmental changes will require additional research that builds on the current findings but also tries to examine in more detail the role of learning strategy and process.

Chapter 5

Scaling up

> "[Learners] *are like sailors who on the open sea must reconstruct their ship but are never able to start afresh from the bottom. Where a beam is taken away a new one must at once be put there, and for this the rest of the ship is used as support. In this way, by using the old beams and driftwood the ship can be shaped entirely anew, but only by gradual reconstruction.*"
>
> — WILLARD V. O. QUINE

Models of human causal learning based on Bayesian networks have tended to focus on what Marr (1982) called the *computational* level. This means that they consider the abstract computational problem being solved and its ideal solution rather than the actual cognitive processes involved in reaching that solution — Marr's *algorithmic* level. In practice the demands of computing and storing the quantities required for exactly solving the problem of causal learning are intractable for any non-trivial world and plausibly-bounded learner. Even a small number of potential relata permit massive numbers of patterns of causal relationships. Moreover, real learning contexts involve noisy (unreliable) relationships and the threat of exogenous interference, further compounding the complexity of normative inference. Navigating this space of possibilities optimally would require maintaining probability distributions across many models and updating all these probabilities whenever integrating new evidence. This evidence might in turn be gathered piecemeal over a lifetime of experience. Doing so efficiently would require choosing maximally informative interventions, a task which poses even greater computational challenges: consideration and weighting of all possible outcomes, under all possible models for all possible interventions (Murphy, 2001; Nyberg & Korb, 2006).

In order to understand better the cognitive processes involved in learning causal relationships, we present a detailed exploration of how people, with their limited processing resources, represent and reason about causal structure. We begin by surveying existing proposals in the literature. We then draw on the literature on algorithms for approximating probabilistic inference in computer science using these to construct a new model. We show that our new model captures the behavioural patterns using a scalable and cognitively plausible algorithm and explains why aggregate behaviour appears noisily normative in the face of individual heterogeneity.

Many existing experiments on human causal learning involve small numbers of possible structures, semi-deterministic relationships and limited choices or opportunities to intervene. These constraints limit the computational demands on learners, and thus the need for heuristics or approximations. Further, in most existing studies, subjects make causal judgments only at the end of a period of learning, limiting what we can learn about how their beliefs evolved as they observed more evidence, and how this relates to intervention choice dynamics. One exception is Bramley, Lagnado, and Speekenbrink (2015) (Chapter 3), which explored online causal learning in scenarios where participants' judgments about an underlying causal structure were repeatedly elicited over a sequence of interventional tests. Another is Bramley, Dayan, and Lagnado (2015), which built on this paradigm. Both papers explained participants' judgments with accounts that are not completely satisfying algorithmically, lacking cognitively plausible or scalable procedures that could capture the ways in which judgments and intervention choices deviated from the rational norms. Here, we develop the algorithmic level account and demonstrate that it outperforms or equals competitors in modelling the data from both previous papers and a new experiment.

The resulting class of algorithms embodies an old idea about theory change known as the Duhem–Quine thesis (Duhem, 1991). The idea can illustrated by a simile, originally attributed to Otto Van Neurath (1932) but popularised by Quine, in the eponymous quotation at the start of this chapter. The *Neurath's ship* metaphor describes the piecemeal growth and evolution of scientific theories over the course of history. In the metaphor, the theorist (sailor) is cast as relying on their existing theory (ship) to stay afloat, without the privilege of a dry-dock in which to make major improvements. Unable to step back and consider all possible alternatives, the theorist is limited to building on the existing theory, making a series of small changes with the goal of improving the fit.

We argue that people are in a similar position when it comes to their beliefs about the causal structure of the world. We propose that a learner normally maintains only a single

hypothesis about the global causal model, rather than a distribution over all possibilities. They update their hypothesis by making local changes (e.g. adding, removing and reversing individual connections, nodes or subgraphs) while depending on the rest of the model as a basis. We show that by doing this, the learner can end up with a relatively accurate causal model without ever representing the whole hypothesis space or storing all the old evidence, but that their causal beliefs will exhibit a particular pattern of sequential dependence. We provide a related account of bounded intervention selection, based on the idea that learners adapt to their own learning limitations when choosing what evidence to gather next, attempting to resolve local rather than global uncertainty. Together, our *Neurath's ship* model and local-uncertainty-based schema for intervention selection provide a step towards an explanation of how people might achieve a *resource rational* (Griffiths et al., 2015; Simon, 1982) trade-off between accuracy and the cognitive costs of maintaining an accurate causal model of the world.

This chapter is organised as follows. It first highlights the ways in which past experiments have shown human learning to diverge from the predictions of the idealised account detailed in 3.1, using these to motivate two causal judgment heuristics proposed in the literature: *simple endorsement* (Bramley, Lagnado, & Speekenbrink, 2015; Fernbach & Sloman, 2009) and *win-stay, lose-sample* (Bonawitz et al., 2014) before developing a new *Neurath's ship* framework for belief change and active learning. It next shows that participants' overall patterns of judgments and intervention choices are in line with the predictions of this framework across a variety of problems varying in terms of the complexity and noise in the true generative model, and whether the participants' are trained or must infer the noise.

The models are then compared at the individual level, and we find that all three causal-judgment proposals substantially outperform baseline and computational level competitors. While our Neurath's ship provides the best overall fit, we find considerable diversity of strategies across participants. In particular, we find that the *simple endorsement* heuristic emerges as a strong competitor. Additional details about the formal framework and model specification are provided in Appendix A. Also, where indicated, additional figures are provided in Supplementary materials at `http://www.ucl.ac.uk/lagnado-lab/el/nbt`.

5.1 Behavioural patterns and existing explanations

Bramley, Lagnado, and Speekenbrink (2015) (Chapter 3) found that participants' judgments in a sequential active causal learning task resembled probability matching when lumped together, but that individuals' trajectories were not well captured by simply adding decision noise to the Bayesian predictions. Individuals' sequences of judgments were sequentially dependent, or "sticky", compared to the Bayesian predictions, tending to remain the same or similar over multiple elicitations as the objectively most likely structure shifted. At the same time, when participants did change their judgments, they tended to do so in ways that were consistent with the most recently gathered evidence, neglecting evidence gathered earlier in learning. The result was a dual pattern of recency in terms of judgments' consistency with the evidence, and stickiness in terms of consistency with the previous judgments. They found that they could capture these patterns with the addition of two parameters to the Bayesian model. The first was a forgetting parameter, encoding trial-by-trial leakage of information from the posterior as it became the prior for the next test. The second was a conservatism parameter, encoding a non-normatively high probability assigned to the latest causal hypothesis. While the resulting model captured participants' choices, it still made the implausible assumption that learners maintained weighted probabilistic beliefs across the whole hypothesis space and performed efficient active learning with respect to these.

As with Bramley, Lagnado, and Speekenbrink (2015) (Chapter 3), Bonawitz et al. (2014) found that children and adults' online structure judgments exhibited sequential dependence. To account for this they proposed an account of how causal learners might rationally reduce the computational effort of continually reconsidering their model. In their "win-stay, lose-sample" scheme they suggest that learners maintain a single structural hypothesis, only resampling a new hypothesis from the posterior when they see something surprising under their current model, concretely, with a probability that increases as the most recent observation becomes less probable. This scheme guarantees that the learner's latest hypothesis is a sample from the posterior distribution at every point, but does not require them to resample with every new trial. While it captures the intuitive idea that people will tend to stick with a hypothesis until it fails to perform, "win-stay, lose-sample" still requires the learner to store all the past evidence to use when resampling, and does not provide a recipe for how the samples are drawn.[1]

[1]The authors mention that MCMC could be used to draw these samples without representing the full posterior.

Another approach to understanding deviations between people's causal judgments and rational norms comes from the idea that people construct causal models in a modular or piecewise way. For example, Waldmann et al. (2008) propose a *minimal rational model* under which learners infer the relationships between each pair of variables separately without worrying about the dependencies between them, ending up with a modular causal model that allows for good local inferences but which leads to so-called "Markov violations" in more complex inferences where participants fail to respect the conditional dependencies and independences implied by the global model (Rehder, 2014). They show that this minimal model is sufficient to capture participants' judgment patterns in two case studies. Building on this idea of locality, Fernbach and Sloman (2009) asked participants to make judgments following observation of several preselected interventions. They found that participants were particularly bad at inferring Chains, often inferring spurious additional links from the root to the sink node (e.g. $X_1 \rightarrow X_3$ as well as $X_1 \rightarrow X_2$ and $X_2 \rightarrow X_3$), a pattern also observed in Bramley, Lagnado, and Speekenbrink (2015). Fernbach and Sloman proposed that this was a consequence of participants inferring causal relationships through local rather than global computations. In the example, the interventions on X_1 would normally lead to activations of X_3 due to the indirect connection via X_2. If learners attended only to X_1 and X_3 there would be the appearance of a direct relationship. They found that they could better model participants by assuming they inferred each causal link separately while ignoring the rest of the model. Embodying this principle, Bramley, Lagnado, and Speekenbrink (2015) proposed a *simple endorsement* heuristic for online causal learning that would tend to add direct edges to a model between intervened-on variables and any variables that activated as a result, removing edges going to any variables that didn't activate. By doing this after each new piece of evidence, the model exhibited recency as the older edges would tend to be overwritten by newly inferred ones, as well as as capturing the pattern of adding unnecessary direct connections in causal chains. The model did a good job of predicting participants' patterns but was outperformed by the Bayesian model bounded with forgetting and conservatism. Additionally, like any heuristic, *simple endorsement*'s success is conditional on its match to the situation. For instance, *simple endorsement* does badly in cases where there are many chains — meaning that the outcome of many interventions are indirect, and also if the true w_B is high.

Going beyond causal learning, sequential effects are ubiquitous in cognition. In some instances they can be rational; for instance moderate recency is rational in a changing world (Julier & Uhlmann, 1997). Regardless, there are a plethora of non-Bayesian models

that can reproduce various sequential effects (DeCarlo, 1992; Gilden, 2001; Treisman & Williams, 1984). A common class of these is based on the idea of adjusting an estimate part way toward new evidence (e.g. Einhorn & Hogarth, 1986; Petrov & Anderson, 2005; Rescorla & Wagner, 1972). Updating point estimates means that a learner need not keep all the evidence in memory but can instead make use of the location of the point(s) as a proxy for what was learned in the past. Bramley, Dayan, and Lagnado (2015) propose a model inspired by these ideas, that maintains a single hypothesis, but simultaneously attempts to minimise edits along with the number of variables' latest states that the current model fails to explain. The result is a model where the current belief acts as an anchor and the learner tries to explain the latest evidence by making the minimal number of changes. Again, this model provided a good fit with participants' judgments, but did not provide a procedure for how participants were able to search the hypothesis space for the causal structure that minimised these constraints.

In summary, a number of ideas and models have been proposed in the causal and active learning literatures. By design, they all do a good job of capturing patterns in human causal judgments. However, it is not clear that any of these proposals provide a general purpose, scalable explanation for human success in learning a complex causal world-model. Some (e.g. *win-stay, lose-sample*) capture behavioural patterns within the normative framework, but do not provide a scalable algorithm. Others (e.g. *simple endorsement*) provide simple scalable heuristics but may not generalise beyond the tasks they were designed for, nor explain human successes in harder problems. In the next section we take inspiration from methods for approximate inference in machine learning to construct a general purpose algorithm for incremental structure change that satisfies both these desiderata.

5.2 Algorithms for causal learning with limited resources

We now turn to algorithms in machine learning that make approximate learning efficient in otherwise intractable circumstances. Additionally, research in these fields on active learning and optimal experiment design has identified a range of reasonable heuristics for selecting queries when the full expected information calculation (Equation 3.10) is intractable. We will take inspiration from some of these ideas to give a formal basis to the intuitions behind the *Neurath's ship* metaphor. We will then use this formal model to generate predictions that we will compare to participants' behaviour in several experiments.

5.2.1 Approximating with a few hypotheses

One common approximation, for situations where a posterior cannot be evaluated in closed form, is to maintain a manageable number of individual hypotheses, or "particles" (Liu & Chen, 1998), with weights corresponding to their relative likelihoods. The ensemble of particles then acts as an approximation to the desired distribution. Sophisticated reweighting and resampling schemes can then filter the ensemble as data are observed, approximating Bayesian inference.

These "particle filtering" methods have been used to explain how humans and other animals might approximate the solutions to complex problems of probabilistic inference. In associative learning (Courville & Daw, 2007), categorisation (Sanborn et al., 2010) and binary decision making (Vul et al., 2009), it has been proposed that people's beliefs actually behave most like a single particle, capturing why individuals often exhibit fluctuating and sub-optimal judgment while maintaining a connection to Bayesian inference, particularly at the population level.

5.2.2 Sequential local search

The idea that people's causal theories are like particles requires they also have some procedure for sampling or adapting these theories as evidence is observed. Another class of useful machine learning methods involves generating sequences of hypotheses, each linked to the next via a form of possibly stochastic transition mechanism. Two members of this class are particularly popular in the present context: Markov Chain Monte Carlo (MCMC) sampling, which asymptotically approximates the posterior distribution; and (stochastic) hill climbing, which merely tries to find hypotheses that have high posterior probabilities.

MCMC algorithms involve stochastic transitions with samples that are typically easy to generate. Under various conditions, this implies that the sequences of (dependent) sample hypotheses form a Markov chain with a stationary distribution that is the full, intended, posterior distribution (Metropolis et al., 1953). The samples will appear to "walk" randomly around the space of possibilities, tending to visit more probable hypotheses more frequently. If samples are extracted from the sequence after a sufficiently long initial, so-called burn-in, period, and sufficiently far apart (to reduce the effect of dependence), they can provide a good approximation to the true posterior distribution.

There are typically many different classes of Markov chain transitions that share the same stationary distribution, but differ in the properties of burn-in and subsampling.

The stochasticity inherent in MCMC algorithms implies that the sequence sometimes makes a transition from a more probable to a less probable hypothesis — this is necessary to sample multi-modal posterior distributions. A more radical heuristic is only to allow transitions to more probable hypotheses — this is called "hill-climbing", attempting to find, and then stick at, the best hypothesis (Tsamardinos, Brown, & Aliferis, 2006). This is typically faster than a full MCMC algorithm to find a good hypothesis, but is prone to become stuck in a local optimum, where the current hypothesis is more likely than all its neighbours, but less likely than some other more distant hypothesis.

Applied to causal structure inference, we might in either case consider transitions that change at most a single edge in the model (Cooper & Herskovits, 1992; Goudie & Mukherjee, 2011). A simple case is Gibbs sampling (Geman & Geman, 1984), starting with some structural hypothesis and repeatedly selecting an edge (randomly or systematically) and re-sampling it (either adding, removing or reversing) conditional on the state of the other edges. This means that a learner can search for a new hypothesis by making local changes to their current hypothesis, reconsidering each of the edges in turn, conditioning on the state of the others without ever enumerating all the possibilities. By constructing a short chain of such "rethinks" a learner can easily update a singular hypothesis without starting from scratch. The longer the chain, the less dependent or "local" the new hypothesis will be to the starting point.

The idea that stochastic local search plays an important role in cognition has some precedent (Gershman et al., 2012; Sanborn et al., 2010). For instance, Abbott, Austerweil, and Griffiths (2012) propose a random local search model of memory retrieval and Ullman, Goodman, and Tenenbaum (2012) propose an MCMC search model for capturing how children search large combinatorial theory spaces when learning intuitive physical theories like taxonomy and magnetism. The idea that people might update their judgments by something like MCMC sampling is also explored by Lieder, Griffiths and Goodman (2012; under review). They argue that under reasonable assumptions about the costs of resampling and need for accuracy, it can be rational to update one's beliefs by constructing short chains where the updated judgment retains some dependence on its starting state, arguing that this might explain *anchoring* effects (Kahneman et al., 1982).

In addition to computational savings, updating beliefs by local search can be desirable for statistical reasons. If the learner has forgotten some of the evidence they have seen, the location of their previous hypothesis acts like a very approximate version of a sufficient statistic for the forgotten information. This can make it advantageous to the learner to strike a good balance between editing their model to account better for the data they can remember, and staying close to their previous model to retain the connection to the data they have forgotten (Bramley, Lagnado, & Speekenbrink, 2015).

5.3 Neurath's ship: An algorithmic-level model of sequential belief change

The previous section summarised two ideas derived from computer science and statistics that provide a potential solution to the computational challenges of causal learning: maintaining only a single hypothesis at a time, and exploring new hypotheses using local search based on sampling. In this section, we formalise these ideas to define a class of models of causal learning inspired by the metaphor of Neurath's ship. We start by treating interventions as given, and only focus on inference. We then consider the nature of the interventions.

Concretely, we propose that causal learners maintain only a single causal model (a single particle) b^{t-1}, and a collection of recent evidence and interventions $\mathcal{D}_r^{t-1}$ and $\mathcal{C}_r^{t-1}$ at time $t-1$. They then make inferences by:

1. Observing the latest evidence $\mathbf{d}^t$ and $\mathbf{c}^t$ and adding it to the collection to make $\mathcal{D}_r^t$ and $\mathcal{C}_r^t$.

2. Then, searching for local improvements to b^{t-1} by sequentially reconsidering edges $E_{ij} \in \{1 : i \to j,\ 0 : i \nrightarrow j,\ -1 : i \leftarrow j\}$ (adding, subtracting or reorienting them) conditional on the current state of the edges in the rest of their model $E_{\setminus ij}$ — e.g. with probability $P(E_{ij}|E_{\setminus ij}, \mathcal{D}_r^t, \mathcal{C}_r^t, \mathbf{w})$.

3. After searching for k steps, stopping and taking the latest version of their model as their new belief b^t. If b^t differs from b^{t-1} the evidence is forgotten ($\mathcal{D}_r^t$ and $\mathcal{C}_r^t$ become $\{\}$), and they begin collecting evidence again.

A detailed specification of this process is given in Appendix A.

Starting with any hypothesis and repeatedly resampling edges conditional on the others is a form of Gibbs sampling (Goudie & Mukherjee, 2011). Further, the learner can make use of the data they have forgotten by starting the search with their current belief b^{t-1}, since these data are represented to some degree in the location of b^{t-1}. Resampling using the recent data $P(M|\mathcal{D}_r^t, \mathcal{C}_r^t, \mathbf{w})$ allows the learner to adjust their beliefs to encapsulate better the data they have just seen, and let this evidence fall out of memory once it has been incorporated into the model.

5.3.1 Resampling, hill climbing or random change

Following the procedure outlined above, the learner's search steps would constitute dependent samples from the posterior over structures given $\mathcal{D}_r^t$. However, it is also plausible that learners will try to hill-climb rather than sample, preferring to move to more probable local models more strongly than would be predicted by Gibbs sampling. In order to explore this, we will consider generalisations of the update equation allowing transitions to be governed by powers of the conditional edge probability (i.e. $P^\omega(E_{ij} = e|E_{\setminus ij}, \mathcal{D}_r^t, \mathcal{C}_r^t, \mathbf{w})$), yielding stronger or weaker preference for the most likely state of E_{ij} depending whether $\omega > 1$ or < 1. By setting ω to zero, we would get a model that does not learn but just moves randomly between hypotheses, tending to remain local and by setting it to infinity we would get a model that always moved to the most likely state for the edge.

5.3.2 Search length

It is reasonable to assume that the number of search steps k that a learner performs will be variable, but that their *capacity* to search will be relatively stable. Therefore, we assume that for each update, the learner searches for k steps, where k is drawn from a Poisson distribution with mean $\lambda \in [0, \infty]$.

The value of λ thus determines how sequentially dependent a learner's sequences of beliefs are. A large λ codifies a tendency to move beliefs a long way to account for the latest data $\mathcal{D}_r^t$ at the expense of the older data — retained only in the location of the previous belief b^{t-1} — while a moderate λ captures a reasonable trade-off between starting state and new evidence, and a small λ captures *conservatism*, i.e. failure to shift beliefs enough to account for the latest data.[2]

[2]Note that we later cap k at 50 when estimating our model having established that search lengths beyond these bounds made negligible difference to predictions.

5.3.3 Putting these together

By representing the transition probabilities from model i to model j, for a particular setting of hill climbing parameter ω and data $\mathcal{D}_r^t$, with a transition matrix R_t^ω, we can thus make probabilistic predictions about a learner's new belief $b^t \in B^t$.[3] The probabilities depend on the previous belief b^{t-1} and their average search length λ. By averaging over different search lengths with their probability controlled by λ, and taking the requisite row of the resulting transition matrix we get the following equation

$$P(b^t = m|\mathcal{D}_r^t, \mathcal{C}_r^t, b^{t-1}, \omega, \lambda) = \sum_0^\infty \frac{\lambda^k e^{-\lambda}}{k!} [(R_t^\omega)^k]_{b^{t-1}m} \tag{5.1}$$

Note that this equation describes the probability of a Neurath's ship style search terminating in a given new location. The learner themselves need only follow the four steps described above, *sampling* particular edges and search length rather than averaging over the possible values of these quantities. See Appendix A for more details and Figure 5.1 for an example.

5.4 Selecting interventions on Neurath's ship: A local uncertainty schema

In situations where a posterior is already hard to evaluate, calculating the globally most informative intervention — finding the intervention $\mathbf{c}^t$ that maximises Equation 3.10 — will almost always be infeasible. Therefore, a variety of heuristics have been developed that allow tests to be selected that are more useful than random selection, but do not require the full expected information gain be computed (Settles, 2012). These tend to rely on the learner's current, rather than expected, uncertainty (e.g. *uncertainty sampling* which chooses based on outcome uncertainty under the prior) or the predictions under just a few favoured hypotheses (e.g. *query by committee*) as a substitute for the full expectancy calculation. The former relies on maintaining a complete prior distribution, making the latter a more natural partner to the *Neurath's ship* framework.

We have proposed a model of structure inference under which learners are only able to consider a small set of alternatives at a time, and only able to generate alternatives

[3] We define this matrix formally in Appendix A. Note that we assume transitions that would create a loop in the overall model get a probability of zero. This assumption could be dropped for learning dynamic Bayesian networks but is necessary for working with directed acyclic graphs.

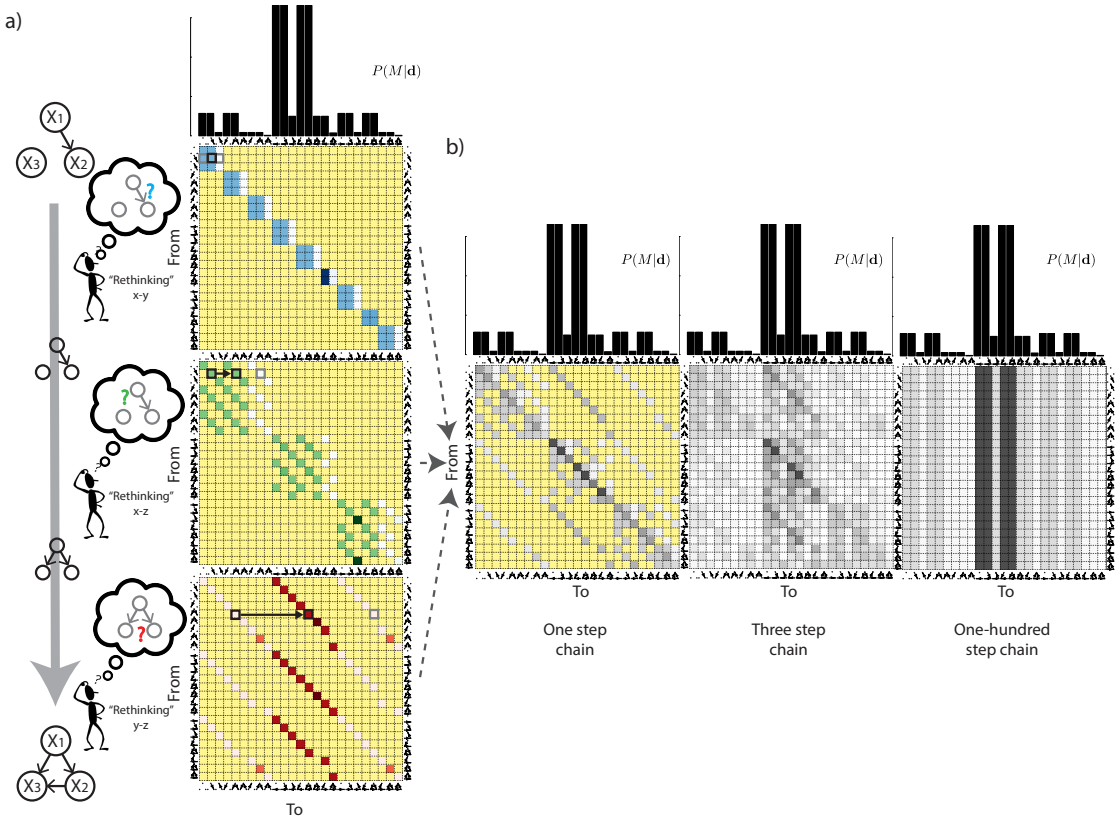


FIGURE 5.1: An illustration of NS model of causal belief updating. a) An example search path: The learner starts out with a singly connected model at the top ($X_1 \to X_2$ connection only). They update their beliefs by resampling one edge at a time $e \in \{\to, \nleftrightarrow, \leftarrow\}$. Each entry i, j in the matrices gives the probability of moving from model in the row i to the model in the column j when resampling the edge marked with the coloured question mark. Lighter shades of the requisite colour indicate low transition probability, darker shades indicate greater transition probability; yellow is used to indicate zero probabilities. Here the learner stops after resampling each edge once, moving from b^{t-1} of $[X_1 \to X_2]$ to b^t of $[X_1 \to X_2, X_1 \to X_3, X_2 \to X_3]$. b) Assuming the edge to resample is chosen at random, we can average over the different possible edge choices to derive a 1-step Markov chain transition matrix R_t^ω encompassing all the possibilities. By raising this matrix to higher powers we get the probability of different end points for searches of that length. If the chain is short (small k) the final state depends heavily on the starting state (left) but for longer chains (large k), the starting state becomes less important, getting increasingly close to independent sampling from the desired distribution (right).

that are "local" in some dimension. Locally driven intervention selection is a natural partner to this for at least two reasons: (1) Under the constraints of the *Neurath's ship* framework, learners would not be able to work with the prospective distributions required to estimate global expected informativeness, but could potentially estimate expected informativeness with respect to a sufficiently narrow sets of alternatives. (2) Evidence optimised to distinguishing local possibilities (focused on one edge at a time for instance) might better support sequential local belief updates (of the kind emphasised in our framework) than the globally most informative evidence (Patil, Zhu, Kopeć, &

Love, 2014). In line with this, we propose one way in which learners might select robustly informative interventions, by attempting only to distinguish a few "local" possibilities at a time, requiring only "local" uncertainty estimates to target the possibilities on which to focus (Markant et al., 2015).

The idea that learners will focus on distinguishing only a few alternatives at a time requires specifying how they choose which of the many possible subsets of the full hypothesis space to target with a particular test. Queries that optimally reduce expected uncertainty about one local aspect of a problem are liable to differ from those that promise high global uncertainty reduction. For example, Figure 5.2b shows two trials taken from our experiments, and shows that the expected values of each of a range of different intervention choices (shown in Figure 5.2a) are very different depending on whether the learner is focused on resolving global uncertainty all at once, or on resolving some specific "local" aspect of it. This illustrates the idea that a learner might choose a test that is optimally informative with respect to a modest range of options that they have in mind at the time (e.g. models that differ just in terms of the state of E_{xz}) yet appear sporadically inefficient from the perspective of greedy global uncertainty reduction. Furthermore, by licensing quite different intervention preferences, they allow us to diagnose individual and trial-by-trial differences in focus preference.

In the current work, we will consider three possible varieties of focus, one motivated by the *Neurath's ship* framework (*edge* focus) and two inspired by existing ideas about bounded search and discovery in the literature (*effects* focus and *confirmation* focus). While these are by no means exhaustive they represent a reasonable starting point.

5.4.1 The two stages of the schema

The idea that learners focus on resolving local rather than global uncertainty results in a metaproblem of choosing what to focus on next, making intervention choice a two stage process. We write L for the set of all possible foci $\mathbf{l}$, and $\mathcal{L} \subset L$ for the subset of possibilities that the learner will consider at a time, such as the the state of a particular edge or the effects of a particular variable. The procedure is:

Stage 1 Selecting a local focus $\mathbf{l}^t \in \mathcal{L}$

Stage 2 Selecting an informative test $\mathbf{c}^t$ with respect to the chosen focus $\mathbf{l}^t$

Different learners might differ in the types of questions they consider, meaning that $\mathcal{L}$ might contain different varieties and combinations of local focuses. We first formalise the two stages of the schema, and then propose three varieties of local focus that learners might consider in their option set $\mathcal{L}$ that differ in terms of which and how many alternatives they include.

As mentioned above, we assume that the learner has some way of estimating their current local confidence. We will assume confidence here is approximately the inverse of uncertainty, so assume for simplicity that learners can calculate uncertainty from the evidence they have gathered since last changing their model in the form of the entropy $H(\mathbf{l}|\mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t)$ for all $\mathbf{l} \in \mathcal{L}$ (the assumption we examine in the discussion). They then choose (**Stage 1**) the locale where these data imply the least certainty

$$\mathbf{l}^t = \arg\max_{\mathbf{l}\in\mathcal{L}} H(\mathbf{l}|b^{t-1}, \mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t) \tag{5.2}$$

However, in carrying out **Stage 2** we make the radical assumption that learners do not use $P(\mathbf{l}^t|\mathcal{D}_r^t, b^{t-1}, \mathbf{w}; \mathcal{C}_r^t)$, but rather, consistent with the method of inference itself, only consider the potential next datum $\mathbf{d}'$. This means that the intervention $\mathbf{c}^t$ itself is chosen to maximise the expected information about $\mathbf{l}^t$, ignoring pre-existing evidence, and using what amounts to a uniform prior. Specifically, we assume that $\mathbf{c}^t$ is chosen as

$$\mathbf{c}^t = \arg\max_{\mathbf{c}\in\mathcal{C}} \mathop{\mathbb{E}}_{\mathbf{d}\in\mathcal{D}_\mathbf{c}} \left[\Delta H(\mathbf{l}^t|\mathbf{d}, \mathbf{w}, b^{t-1}; \mathbf{c})\right] \tag{5.3}$$

where we detail the term in the expectation below for the three types of focuses.

Assuming real learners will exhibit some decision noise, we can model both choice of focus and choice of intervention relative to a focus as soft (Luce, 1959) rather than strict maximisation giving focus probabilities

$$P(\mathbf{l}^t|\mathcal{D}_r^t, b^{t-1}, \mathbf{w}; \mathcal{C}_r^t) = \frac{\exp(H(\mathbf{l}^t|\mathcal{D}_r^t, b^{t-1}, \mathbf{w}; \mathcal{C}_r^t)\rho)}{\sum_{\mathbf{l}\in\mathcal{L}} \exp(H(\mathbf{l}|\mathcal{D}_r^t, b^{t-1}, \mathbf{w}; \mathcal{C}_r^t)\rho)} \tag{5.4}$$

governed by some inverse temperature parameter ρ, and choice probabilities

$$P(\mathbf{c}^t|\mathbf{l}, \mathbf{w}, b^{t-1}) = \frac{\exp(\mathbb{E}_{\mathbf{d}'\in\mathcal{D}_\mathbf{c}} \left[\Delta H(\mathbf{l}|\mathbf{d}', \mathbf{w}, b^{t-1}; \mathbf{c}^t)\right] \eta)}{\sum_{c\in\mathcal{C}} \exp(\mathbb{E}_{\mathbf{d}'\in\mathcal{D}_\mathbf{c}} \left[\Delta H(\mathbf{l}|\mathbf{d}', \mathbf{w}, b^{t-1}; \mathbf{c})\right] \eta)} \tag{5.5}$$

governed by an inverse temperature η.

5.4.2 Three varieties of local focus

Edges

An obvious choice, given the *Neurath's ship* framework, would be for learners to try to distinguish alternatives that differ in terms of a single edge (Figure 5.2a), i.e. those they would consider during a single update step.

For a chosen edge E_{xy} we can then consider a learner's goal to be to maximise their expectation of

$$\Delta H(E_{xy}|E^{t-1}_{\setminus xy}, \mathbf{d}, \mathbf{w}; \mathbf{c}) \tag{5.6}$$

(see Appendix A for the full local entropy equations). Note that Equation 5.6 is a refinement of Equation 5.3 for the case of focusing on an edge, from b^{t-1} the learner need only condition on the other edges $E^{t-1}_{\setminus xy}$. This goal results in a preference for fixing one of the nodes of the target edge "on", leaving the other free, and depending on the other connections in b^{t-1}, either favours fixing the other variables "off" or is indifferent about whether they are "on", "off" or "free" (Figure 5.2b). For an edge focused local learner, the set of possible focuses includes all the edges $\mathcal{L} \in \forall_{i<j\in N} E_{ij}$.

Effects

A commonly proposed heuristic for efficient search in the deterministic domains is to ask about the dimension that best divides the hypothesis space, eliminating the greatest possible number of options on average. This is variously known as "constraint-seeking" (Ruggeri & Lombrozo, 2014) or "the split half heuristic" (Nelson et al., 2014). In the case of identifying the true deterministic ($w_S = 1$ and $w_B = 0$) causal model on N variables through interventions it turns out that the best split is achieved by querying the effects of a randomly chosen variable, essentially asking: "What does X_i do?" (Figure 5.2a)[4]. Formally we might think of this question as asking: which other variables (if any) are descendants of variable X_1 in the true model? This a broader focus than querying the state of a single edge, but considerably simpler question than the global "which is the right causal model?" because the possibilities just include the different combinations of the other variables as effects (e.g. *neither*, *either* or *both* of X_2 and X_3

[4]This is also the most globally informative type of test relative to a uniform prior in all of the noise conditions we consider in the current paper

are descendants of X_1 in a 3-variable model) rather than the superexponential number of model possibilities[5].

Relative to a chosen variable X_i, we can write an *effect focus* goal as maximising the expectation of

$$\Delta H(\mathrm{De}(X_i)|\mathbf{d}, \mathbf{w}; \mathbf{c}) \tag{5.7}$$

where $\mathrm{De}(X_i)$ is the set of X_i's direct or indirect descendants. This focus does not depend on b^{t-1}. This goal results in a preference for fixing the target node "on" (e.g. $\mathrm{Do}[X_1 = 1]$) and leaving the rest of the variables free to vary (Figure 5.2b). For an effect focused local learner, the set of possible focuses includes all the nodes $\mathcal{L} \in \forall_{i \in X} \mathrm{De}(X\{i\})$.

Confirmation

Another form of local test, is to seek evidence that would confirm or refute the current hypothesis, against a single alternative "null" hypothesis. Confirmatory evidence gathering is a ubiquitous psychological phenomenon (Klayman & Ha, 1989; Nickerson, 1998). Although confirmation seeking is widely touted as a bias, it can also be shown to be optimal, for example under deterministic or sparse hypotheses spaces or peaked priors (Austerweil & Griffiths, 2011; Navarro & Perfors, 2011).

Accordingly, Coenen et al. (2015) propose that causal learners adopt a "positive test strategy" when distinguishing causal models. They define this as a preference to "turn on" a parent component of one's hypothesis — observing whether the activity propagates to the other variables in the way that this hypothesis predicts. They find that people often intervene on suspected parent components, even when this is uninformative, and do so more often under time pressure. In Coenen et al.'s tasks, the goal was always to distinguish between two hypotheses, so their model assumed people would sum over the number of descendants each variable had under each hypothesis and turn on the component that had the most descendants on average. However, this does not generalise to the current, unrestricted, context where all variables have the same number of descendants if you average over the whole hypothesis space. However, Steyvers et al. (2003) propose a related *rational test model* that selects interventions with a goal of distinguishing a single current hypothesis from a null hypothesis that there is no causal connection.

[5]The number of directed acyclic graphs on N nodes, $|\mathcal{M}|_N$, can be computed with the recurrence relation $|\mathcal{M}|_N = \sum_{k \in N} (-1)^{k-1} \binom{N}{2} 2^{k(N-k)} |\mathcal{M}|_{N-1}$ (see Robinson, 1977)

Following Steyvers et al. (2003), for a *confirmatory* focus we consider interventions expected to best reduce uncertainty between the learner's current hypothesis b^{t-1} and a null b^0 in which there are no connections (Figure 5.2a).

$$\Delta H(\{b^t, b^0\}|b^{t-1}, \mathbf{d}, \mathbf{w}; \mathbf{c}) \tag{5.8}$$

This goal results in a preference for fixing on the root node(s) of the target hypothesis (Figure 5.2c ii, noting the confirmation focus favours $\text{Do}[X_1 = 1, X_2 = 1]$ here). The effectiveness of confirmatory focused testing depends on the level of noise and the prior, becoming increasingly useful later once the model being tested has sufficiently high prior probability. For a confirmation focused learner there is always just a single local focus.

5.4.3 Implications of the schema

The local uncertainty schema implies that intervention choice depends on two separable stages. Thus, it accommodates the idea that a learner might be poor at choosing what to focus on but good at selecting an informative intervention relative to their chosen focus. It also allows that we might understand differences in learners' intervention choices as consequences of the types of local focus they are inclined or able to focus on. Learners cognisant of the limitations in their ability to incorporate new evidence might choose to focus their intervention on narrower questions (i.e. learning about a single edge at a time) while others might focus too broadly and fail to learn effectively. In the current work we will fit behaviour assuming that learners choose between these local focuses, using their patterns to diagnose which local focuses they include in their option set $\mathcal{L}$, which of these they choose on a given test l^t and finally how these choices relate to their final performance.

5.5 Comparing model predictions to experiments

The *Neurath's ship* framework we have introduced has two distinct signatures. Making only local edits from a single hypothesis results in sequential dependence. Making these edits by local resampling leads to aggregate behaviour that can range between probability matching and hill climbing — which can give better short term gains but with a tendency to get stuck in local optima. Two of the other heuristics also lead to sequential dependence. *Win-stay lose-sample* predicts all-or-none dependence whereby learners'

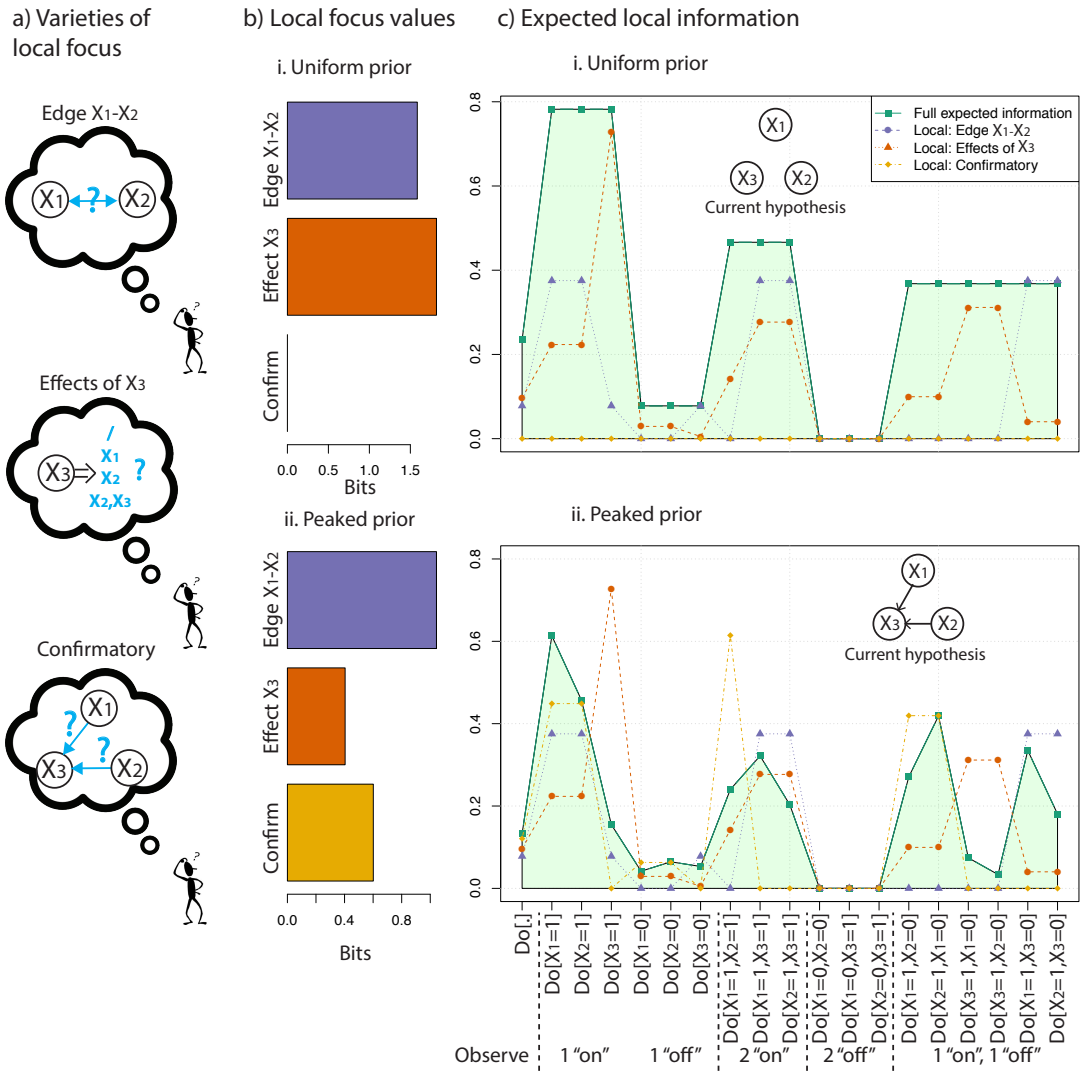


FIGURE 5.2: An illustrative example of local focused uncertainty minimisation a) Three possible "local" focuses. b) The value of these choices of focus according to their current uncertainty (Equation 5.2) i. at the start of learning and ii. after several tests have been performed. Note that uncertainty is measured with Shannon entropy based on the local possibilities and $\mathcal{D}_r^t$ and that confirmation is undefined at the start of learning where both the current and null hypothesis are that there are no connections in the model. c) Expected value of 19 different interventions assuming: global expected information gain from the true prior (green squares, and shaded), effects of X_3 focus (red circles), the relationship between X_1 and X_2 (blue triangles) and confirming b^{t-1} (yellow diamonds), assuming a uniform prior over the requisite possibilities and a known w_S and w_B of .85 and .15.

judgments will either stay the same or jump to a new location that depends only on the posterior. *Simple endorsement* also predicts recency, although it is distinguished by its failure to separate direct from indirect effects of interventions, leading to a different pattern of structural change.

In terms of interventions, if participants are locally focused, we expect their hypotheses to deviate from optimal predictions in ways that can be accommodated by our local

uncertainty schema, i.e. selecting interventions that are more likely to be targeted toward local rather than global uncertainty. If learners do not maintain the full posterior, we expect their intervention distributions to be relatively insensitive to the evidence that has already been seen, while still being locally informative. If people disproportionately focus on identifying *effects*, we expect to see relatively unconstrained interventions with one variable fixed "on" at a time. If people focus on individual *edges* we expect more constraining interventions with more variables fixed "off". If *confirmatory* tests are employed, we expect to see more interventions on putative parents than on child nodes.

We first compare the predictions of our framework to existing data from Bramley, Lagnado, and Speekenbrink (2015). We then report on three new experiments designed to further test the specific predictions of our framework.

5.5.1 Bramley, Lagnado & Speekenbrink (2015)

In Experiments 1 and 2 (Bramley, Lagnado, & Speekenbrink, 2015), participants interacted with five probabilistic causal systems involving 3 variables (see Figure 5.5a), repeatedly selecting interventions (or tests) to perform in which any number of the variables are either fixed "on" or "off", while the remainder are left free to vary. The tests people chose, along with the parameters $\mathbf{w}$ of the true underlying causal model, jointly determined the data they saw. In this experiment w_S was always .8 and w_B was always .1. After each test, participants registered their best guess about the underlying structure. They were incentivised to report their best guess about the structure, through receipt of a bonus for each causal relation (or non-relation) correctly registered at the end. There were three conditions: *no information* (N=79) was run first. After discovering that a significant minority of participants performed at chance, condition *information* (N=30), added a button that participants could hover over and remind themselves of the key instructions during the task (the noise, strengths, the goal) and condition *information + summary* (N=30) additionally provided a visual summary of all previous tests and their outcomes.[6] Participants could draw cyclic causal models if they wanted (e.g. $X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_1$) and were not forced to select something for every edge from the start but instead could leave some or all of the edges as "?". Once a relationship was selected they could not return to "?". The task is available online at `http://www.ucl.ac.uk/lagnado-lab/el/nbt`.

[6] In the paper this was reported as two experiments, the second with two between-subjects conditions. They share identical structure and were subsequently analysed together. Therefore we do the same here, reporting as a single experiment with three between-subjects conditions.

Comparing judgment patterns

We compared participants' performance in Experiments 1 and 2 to that of several simulated learners. *Posterior* draws a new sample from the posterior for each judgment. *Random* simply draws a random graph on each judgment. *Neurath's ship* follows the procedure detailed in the previous section, beginning with its previous judgment (b^{t-1}, or an unconnected model at t=1) and reconsidering one edge at a time based on the evidence gathered since its last change $\mathcal{D}_r^t$ for a small number of steps after observing each outcome. We illustrate this with a simulation with a short mean search length λ of 1.5 and behaviour ω of 10 corresponding moderate hill climbing. *Win-stay, lose-sample* sticks with the previous judgment with probability $1 - P(D^t|b^{t-1}\mathbf{w}; C^t)$ or alternatively samples from the full posterior. The *simple endorser* always adds edges from any intervened-upon variables to any activated variables on each trial, and removes them from any intervened-upon variables to any non-activated variables, overwriting any edges going in the opposing direction. Participants' final accuracy in Experiments 1 and 2 was closest to the *Neurath's ship* as is clear in Figure 5.3a and b. That the Neurath's ship simulation underperformed participants in condition *information + summary* is to be expected since these participants were given a full record of past tests while *Neurath's ship* uses only the recent data.

Additionally, participants' online judgments exhibited sequential dependence. This can be seen in Figure 5.3b comparing the distribution of edits (bars) to the markedly larger shifts we would expect to see assuming random or Bayesian posterior sampling on these trials (black full and dotted lines). The overall pattern of edit distances from judgment to judgment is commensurate with those produced by the *Neurath's ship* procedure (red line), but also here by *win-stay, lose-sample* (blue line) and *simple endorser* (green line) simulations.

Comparing intervention patterns

To compare intervention choices to global and locally driven intervention selection, we simulated the task with the same number of simulations as participants, stochastically generating the outcomes of the simulations' intervention choices according to the true model and true $\mathbf{w}$ (which the participants knew). Simulated efficient active learners would perfectly track the posterior and always select the greediest intervention (as in Equation 3.10).

We also compared participants' interventions to those of several other simulated learners, each restricted to one of the three types of local focus introduced in Section 4 ('edge'. 'effects' or 'confirmation').[7] When one of the simulated learners did not generate a unique best intervention, it would sample uniformly from the joint-best interventions according to that criterion. The results of the simulations are visualised in Figure 5.3c and d.

Participants' intervention choices in Experiments 1 and 2 were clearly more informative than random selection but less so than ideal active learning. This is evident in Figure 5.3c comparing participants (bars) to simulations of ideal active learning (black circles) and random intervening (black squares), and in Figure 5.3d comparing the participants (red lines) to the ideal active learning (pink lines) and random intervening (blue lines) simulations. Furthermore, the informativeness of participants' interventions is in the range of the simulations of any of the three local foci (yellow, green and blue lines).

As we see in Figure 5.3d, idealised active learning favoured fixing one variable on at a time ($\text{Do}[X_1 = 1]$, $\text{Do}[X_2 = 1]$ etc, hereafter called "one-on" interventions) for the majority of tests. It always chose "one-on" for the first few tests but would sometimes select controlled (e.g. $\text{Do}[X_1 = 1, X_2 = 0]$) tests on later tests when the remaining uncertainty was predominantly between direct and indirect causal pathways as in between Chain, Fork and Fully-connected structures.

Locally driven testing had different signatures depending on the focus. The *edge* focused simulation would fix the component at one end of their edge of interest "on" and leave the component at the other end "free". What it did with the third component depended on its latest judgment about the network. If, according to b^{t-1}, another component was a cause of the component that was left free-to-vary, the simulation favoured fixing it "off". Otherwise, it did not distinguish between "on", "off" or "free" choosing one of these at random. The resulting pattern is a spread across "one-on", "two-on" and "one-on, one-off" tests with a bias toward controlled "one-on, one-off" tests. The effects focused learner always favoured "one-on" interventions. The *confirmation* focused tester would generally fix components with children in b^{t-1} on, and leave components with parents in b^{t-1} free. This led to the choice of a mixture of "one-on" and "two-on" interventions.

[7] We assumed these tests were chosen based on a uniform prior over the options considered. We used the latest most probable judgment $\text{argmax}_{\mathbf{x}} p(M|D^{t-1}, \mathbf{w})$ in place of a current hypothesis b^{t-1} for edge focused and confirmatory testing so as not to presuppose a particular belief update rule in assessing intervention selection.

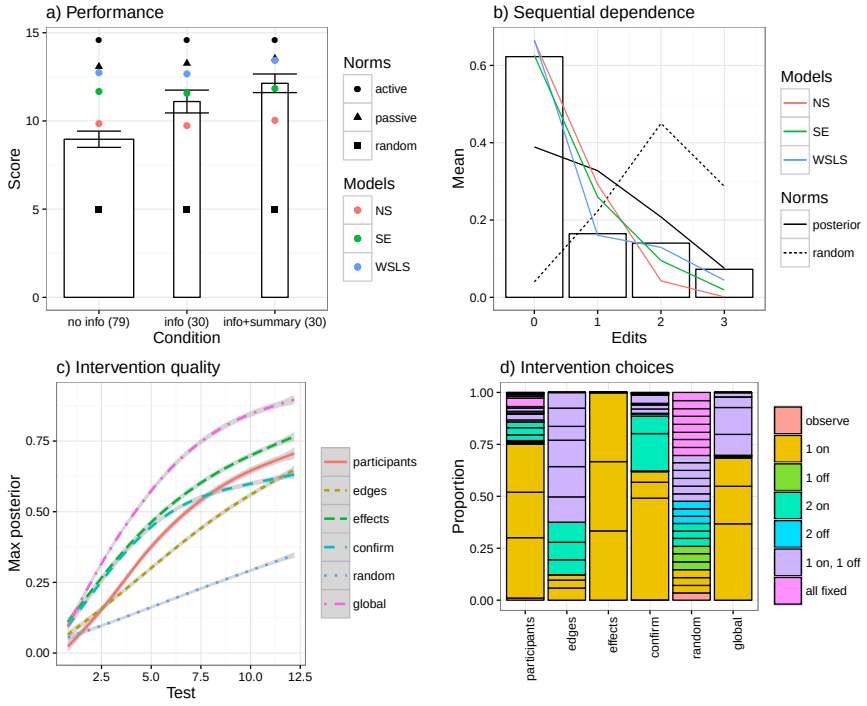


FIGURE 5.3: Experiments 1 and 2; performance and interventions. a) Accuracy by condition. Bars show participant accuracy by condition, and points compare with the models, bar widths visualise the number of participants per condition. b) Sequential dependence. The number of edits made by participants between successive judgments, bars give proportion of participants' updates with different numbers of edits, lines compare with the models. c) Quality of participants' and simulated learners' intervention choices measured by the probability that an ideal learner would guess the correct model given the information generated. The plot shows values smoothed with R's `gam` function and the grey regions give 99% confidence intervals. d) The proportion of interventions of different types chosen by participants as compared to simulated learners. observe = $\text{Do}[\varnothing]$, 1 on = e.g. $\text{Do}[X_1 = 1]$, 1 off = e.g. $\text{Do}[X_1 = 0]$ and so on. All fixed = e.g. $\text{Do}[X_1 = 0, X_2 = 1, X_3 = 0]$.

Like the *ideal* or the *effects* focused simulations, participants in Experiments 1 and 2 strongly favoured "one-on" tests. Consistent with *confirmatory* testing, components with at least one child according to the latest hypothesis b^{t-1} were more likely to be fixed "on" than components believed to have no children (60% compared to 56% of the time $t(24568) = 3.2, p = .001$).[8] Participants' intervention selections were markedly less dynamic across trials than those of the efficient learner. For example, the proportion of single (e.g. $[X_1 = 1]$) interventions decreased only fractionally on later tests, dropping from 78% to 73% from the first to the last test.

[8] We ran the same number of simulated learners as participants in each experiment and condition to facilitate statistical comparison.

5.5.2 Motivating the new experiments

In analysing Experiments 1 and 2, we found patterns of judgments and interventions broadly consistent with our framework. However, the conclusions we can draw from these data alone are somewhat limited. Firstly, the problems participants faced did not strongly delineate our *Neurath's ship* proposal from other proposed approximations, namely the approximate *win-stay, lose-sample* or the heuristic *simple endorsement* which also predicted similar patterns of accuracy and sequential dependence. Similarly, in terms of interventions, participants' strong preference for "one-on" interventions was consistent with local effect-focused testing. However, "one-on" interventions were also the globally most informative choices for the majority of participants' trials, especially early during learning. Thus, we cannot be confident what participants focused on when selecting their interventions.

Methodologically also, several aspects of Experiments 1 and 2 are suboptimal for testing our framework. Participants were allowed to leave edges unspecified and could also draw cyclic models, both of which complicated our analyses. Furthermore, participants had 12 tests on each problem, allowing an idealised learner to approach certainty given the high w_S and low w_B, and for a significant minority of people to perform at ceiling. These choices limit the incentive for participants to be efficient with their interventions. Additionally, participants were only incentivised to be accurate with their final judgment, meaning we cannot be confident that intermediate judgments always represented their best and latest guess about the model. Finally, participants were not forced to update all their edges after each test, meaning that lazy responding could be confused with genuine sequential dependence of beliefs.

Next, we report on two new experiments that build on the paradigm from Experiments 1 and 2, making methodological improvements, while also exploring harder more revealing problems, and eliciting additional measures, all with the goal of better distinguishing our framework from competitors.

Experiment 4 explores learning in more complex problems than in Experiments 1 and 2, with more variables and a range of strengths w_S and levels of background noise w_B, and fewer interventions per problem. The increased complexity and noise provides more space and stronger motivation for the use of approximations and heuristics. Furthermore, the broader range of possible structures and intervention choices increases the discriminability of our framework from alternatives such as *win-stay, lose-sample* and

simple endorsement, while the shorter problems avoid ceiling effects and ensure participants choose interventions carefully. To ensure participants register their best and latest belief at every time point, we also incentivise participants through their accuracy at random time points during learning. To eliminate the possibility of lazy responding biasing results in favour of *Neurath's ship*, we force participants to mark all edges anew after every test without a record of their previous judgment as a guide.

Experiment 5 inherits the methodological improvements, compares two elicitation procedures, and also takes several additional steps. In the previous studies, participants were pretrained on strength w_S and background noise w_B. This will not generally be true; learners will normally have to take into account their uncertainty about these sources of noise during inference. Therefore, Experiment 5 focuses on cases where participants are not pretrained on $\mathbf{w}$. Additionally, our framework makes predictions about participants' problem representation that go beyond how it should manifest in final structure judgments and intervention choices. Specifically, our local intervention schema proposes that people focus on subparts of the overall problem during learning, switching between these by comparing their current local uncertainty. Experiment 5 probes these assumptions by asking learners for confidence judgments about the edges in the model during learning, and eliciting free explanations of what interventions are supposed to be testing. When we go on to fit our framework to individuals in the final section of the paper, we are able to code up these free responses in terms of the hypotheses they refer to and compare them to the focuses predicted by our local uncertainty schema.

5.6 Experiment 4: Learning larger causal models

Our first new experiment looks at learning in harder problems with a range of w_S and w_B and a mixture of 3- and 4-variable problems, asking whether we now see a clearer signature of *Neurath's ship*, *simple endorsement* or *win-stay,lose-sample* style local updating or of local focus during interventions selection.

5.6.1 Methods

Participants

120 participants (68 male, mean±SD age 33±9) were recruited from Amazon Mechanical Turk, split randomly so that 30 performed in each of 4 conditions. They were paid $1.50

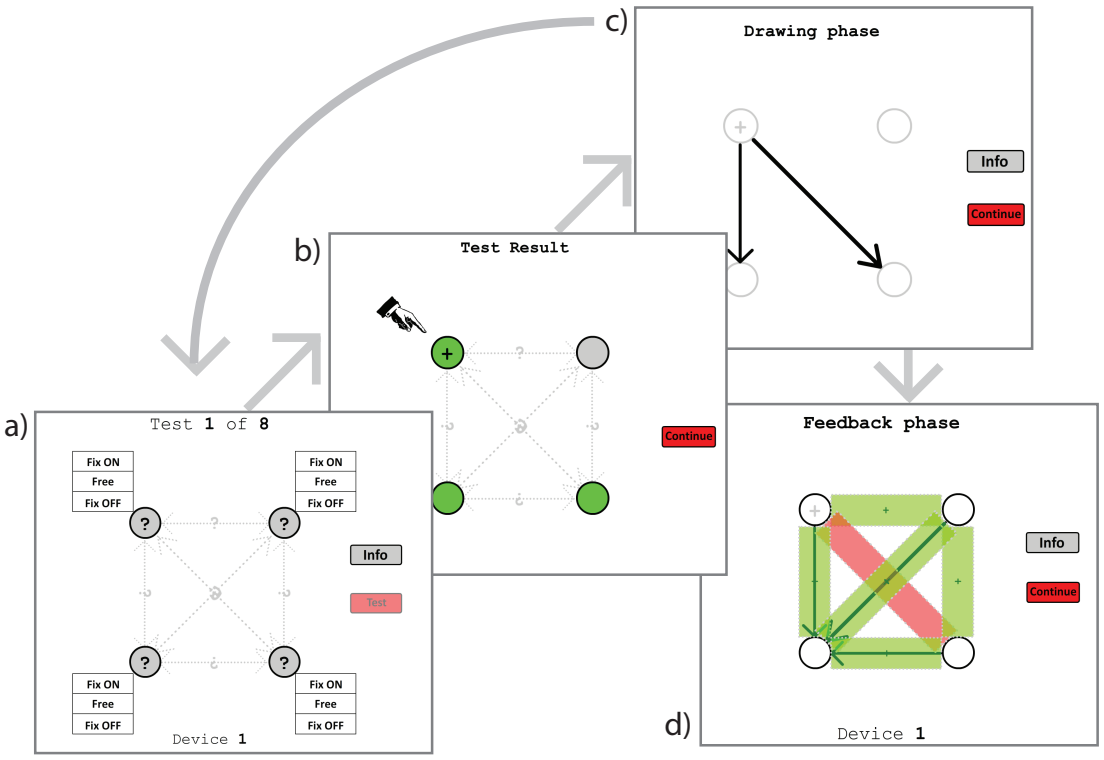


FIGURE 5.4: Experimental procedure. a) Selecting a test b) Observing the outcome c) Updating beliefs d) Getting feedback.

and received a bonus of 10c per correctly identified connection on a randomly chosen test for each problem (max = \$6.00, mean±SD \$3.7 ± 0.65). The task took an average of 44 ± 40 minutes.

Design

This study included the five 3-variable problems in Experiments 1 and 2 plus five additional 4-variable problems (see Figure 5.5a). There were problems exemplifying three key types of causal structure: Colliders (converging connections), Chains (sequential connections) and Forks (diverging connections). Within these, the sparseness of the causal connections varied between a Single connection (devices 1 and 6) and Fully-connected (devices 5 and 10).

There were two different levels of causal strength $w_S \in [.9, 0.75]$ and two different levels of background noise $w_B \in [.1, .25]$ making $2 \times 2 = 4$ between-subjects conditions. For instance, in condition $w_S = .9; w_B = .1$ the causal systems were relatively reliable, with nodes rarely activating without being intervened on, or caused by, an active parent, and connections rarely failing to cause their effects. Meanwhile, in condition $w_S = 0.75; w_B = 0.25$ the outcomes were substantially noisier, with probability .25 that a variable with

no active parent would activate, compared to a probability $1 - (1 - .75)(1 - .25) = 0.81$ for a variable with one active parent.

Procedure

The task interface was similar to that in Experiments 1 and 2. Each device was represented as several grey circles on a white background (see Figure 5.4). Participants were told that the circles were components of a causal system of binary variables, but were not given any further cover story. Initially, all components were inactive and no connection was marked between them. Participants performed tests by clicking on the components, setting them at one of three states "fixed on", "fixed off" and "free-to-vary", then clicking "test" and observing what happened to the "free-to-vary"components as a result. The observations were of temporary activity (graphically, activated components would turn green and wobble).

As in Experiments 1 and 2, participants registered their best guess about the underlying structure after each test. They did this by clicking between the components to select either no connection, or a forward or backward connection (represented as black arrows). Participants were incentivised to be accurate, but unlike in Experiments 1 and 2, payments were based on randomly selected time points rather than the final judgments.

Participants completed instructions familiarising them with the task interface; the interpretation of arrows as (probabilistic) causal connections; the incentives for judgment accuracy. Participants were told background noise level and strength parameters $\mathbf{w}$ explicitly. They were then shown unconnected components and forced to test them several times. The frequency with which the components activated reflected the true background noise level. They were then shown a set of two-component causal systems in which component "X_1" was a cause of "X_2", and were forced to test these systems several times with component X_1 fixed on. This indicated that the frequency with which X_2 activated reflected the level of w_S combined with the background noise they had already learned.

After completing the instructions, participants had to answer four comprehension check questions. If they got any wrong they had to go back to the start of the instructions and try again. Then, participants solved a practice problem randomly drawn from the problem set. They then faced the test problems in random order, with randomly oriented unlabelled components. They performed six tests on each three variable problem, and

eight tests on each four variable problem. After the final test for each problem they received feedback telling them the true connections.

To ensure that participants' judgments were always genuine directed acyclic graphs, participants were told in the instructions that the true causal structure would not contain a loop. Unlike in Experiments 1 and 2, if participants tried to draw a model containing a cyclic structure they would see a message saying "you have drawn connections that make a loop, change or remove one to continue".

As in Experiments 1 and 2 conditions *information* and *information + summary*, participants could hover their mouse over a button for a reminder of the key instructions during the task, but unlike condition *information + summary*, they saw no record of their previous tests and outcomes.

The task can be tried out at `http://www.ucl.ac.uk/lagnado-lab/el/nbt`.

5.6.2 Results and discussion

Judgments

In spite of the considerably greater noise and complexity than Experiments 1 and 2, participants performed significantly above chance in all four conditions (comparing to chance performance of $\frac{1}{3}$, participants' scores differed significantly by t-test with $p < .001$ for all four conditions). They also significantly underperformed a Bayes optimal observer ($p < .001$ for all four conditions, Figure 5.6a). Performance declined as background noise w_B increased $F(1, 118) = 4.3, \eta^2 = .04, p = .04$ but there was no evidence for a relationship with strength w_S $F(1, 118) = 2.7, \eta^2 = .04, p = 0.1$. Judgment accuracy was no lower for four compared to three variable problems $t(238) = 0.76, p = 0.44$. Table 5.1 shows accuracy by device type across all experiments. Accuracy differed by device type $\chi^2 = (4) = 22, p < .001$. Consistent with the idea that people struggle most to distinguish the Chain from the Fork or the Fully-connected model, accuracy was lowest for Chains (devices 3; 8) and second lowest for Fully-connected (5; 10) models.

In all four conditions, participants' final accuracy was closer to that of the *Neurath's ship* simulations than the *simple endorser*, *win-stay, lose sample* or *random* responder or ideal (*passive*) responding (Figure 5.6a).[9]

[9]On the rare occasions where the *simple endorser* procedure would induce a cycle (0.4% of trials), the edges were left in their original state.

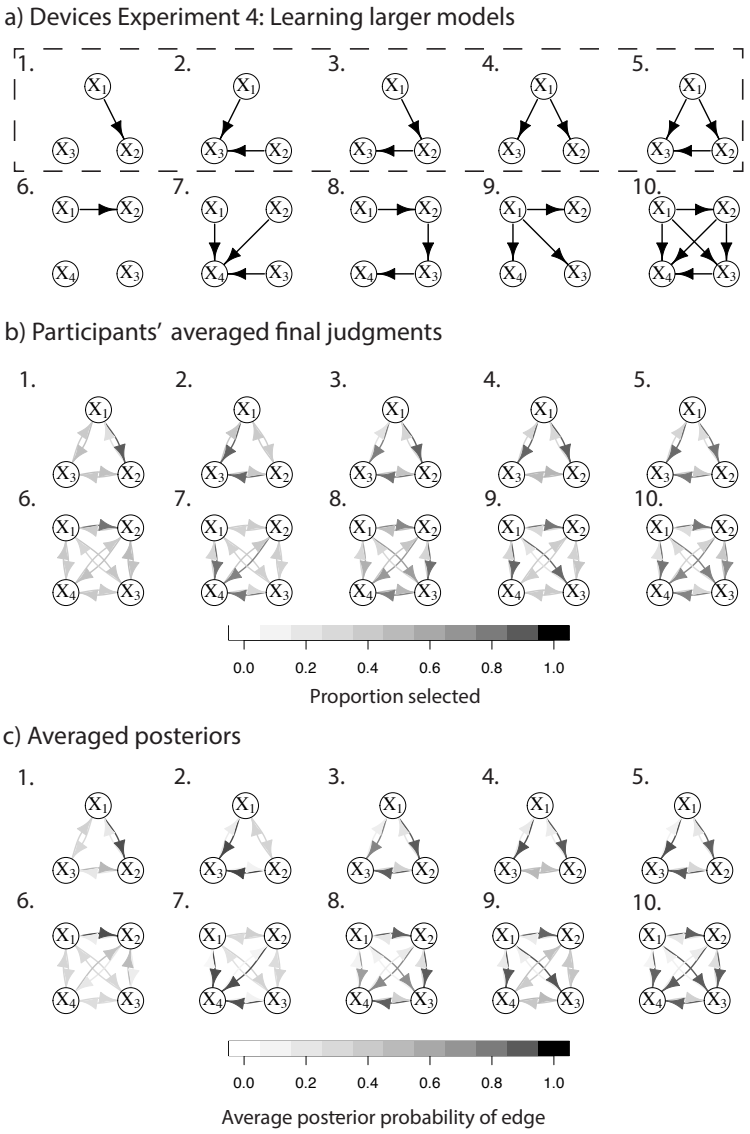


FIGURE 5.5: The true models from Experiment 4: Learning larger models, and visualisation of averaged judgments and posteriors. a) The problems faced by participants. Dashed box indicates those that also appeared in Bramley, Lagnado, and Speekenbrink (2015). b) Averaged final judgments by participants. Darker arrows indicate that a larger proportion of participants marked this edge in their final model. c) Bayes-optimal final marginal probability of each edge in $P(M|D^T, E^T, \mathbf{w})$, averaged over participants' data.

Sequential dependence

Table 5.2 summarises the number of edits (additions, removals or reversals of edges) participants made between each judgment in all experiments. Inspecting the table and Figure 5.6b we see participants' judgments (both high and low performing) show a pattern of rapidly decreasing probability for larger edit distances mimicked by both *Neurath's ship* and *simple endorsement* simulations. In contrast, *random* or *posterior* sampling

Table 5.1: Proportion of Edges Correctly Identified by Device Type in Experiments 1, 2, 4 and 5

Experiment	Variables	Single	Collider	Chain	Fork	Fully-connected	No connection
Exps 1–2	3	0.69	0.71	0.66	0.63	0.67	
Exp 4: Larger models	3	0.57	0.6	0.51	0.56	0.55	
Exp 4: Larger models	4	0.56	0.57	0.48	0.58	0.49	
Exp 5: Unknown strengths	3	0.62	0.6	0.61	0.65	0.61	0.61
All		0.61	0.62	0.57	0.61	0.58	0.61

Table 5.2: Edit Distance Between Consecutive Judgments in Experiments 1, 2, 4 and 5.

Experiment	Var	Participants		Random		SE		WSLS		NS		Posterior	
		M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Exps 1–2	3	0.66	0.97	1.99	0.83	0.51	0.74	0.55	0.88	0.38	0.57	1.02	0.93
Exp 4	3	0.92	1.01	1.99	0.83	0.44	0.64	0.73	0.98	0.50	0.63	1.47	0.94
Exp 4	4	1.69	1.63	3.96	1.16	0.65	0.88	1.64	1.78	0.55	0.68	2.94	1.40
Exp 5 (rem)	3	0.73	0.85	2.02	0.81	0.43	0.62	1.06	1.13	0.59	0.69	1.86	0.85
Exp 5 (dis)	3	1.02	0.99	2.02	0.81	0.43	0.62	1.06	1.13	0.59	0.69	1.86	0.85
All		0.99	1.09	2.40	0.89	0.50	0.71	0.99	1.17	0.51	0.65	1.80	1.00

Note: Var = number of variables. NS = *Neurath's ship* simulations with $\lambda = 1.5$ and $\omega = 10$, WSLS = *win-stay, lose-sample* simulations, SW = *simple Endorser* simulations, M=mean, SD = standard deviation. Rem = remain condition. Dis = disappear condition.

lead to quite different signatures with larger jumps being more probable. Choices simulated from *Neurath's ship* and *simple endorsement* were more sequentially dependent than participants' on average but have the expected decreasing shape. *Win-stay, lose sample* produces a different pattern with a maximum at zero changes but a second peak in the same location as for *posterior sampling* but has an average edit distance very close to that averaged over participants. However, we expect any random or inattentive responding to inflate average edit distances, and indeed find a strong negative correlation between edit distance and score $F(1, 118) = 34, \beta = -6.7, \eta^2 = .34, p < .001$. A simple way to illustrate this is to compare the edits of higher and lower performers. Scores of $\frac{22}{45}$ or more differ significantly from chance performance (around $\frac{15}{45}$) by χ^2 test. The 79 participants that scored 22 or more made markedly smaller edits than those that scored under 22 (0.85 ± 0.95 compared to 1.3 ± 1.12 for three variable, and 1.4 ± 1.5 compared to 2.4 ± 1.8 for four variable problems), putting the clearly successful participants' patterns closer to the "Neurath's ship" and "simple endorser" simulations. Additionally, we expect individual differences in search length λ under the *Neurath's ship* model and here only simulate assuming a mean search length of 1.5. Aggregating over a wider set of simulated learners with different capacities to search for updates would lead to a heavier-tailed distribution of edit distances that would resemble the participants' choices more faithfully.

Interventions

Globally focused active learning favoured a mixture of "one-on" and "one-on, one-off" interventions (and several others including "one-on, two-off" in the four variable problems). The number and nature of the fixed variables it favoured depended strongly on the condition, favouring fixing more variables off when w_S was high. It would also shift dramatically over trials always favouring "one-on" interventions for the first trials but these dropping below 50% of choices by the final test. Participants' choices were much less reactive to condition or trial. There were no clear differences in intervention choices by condition (see supplementary figures available at `http://www.ucl.ac.uk/lagnado-lab/el/nbt`) but participants were a little more likely to select "one on" interventions on their first test (57%) compared to their last test (50%), $t(238) = 1.7, p = .01$. Like the *ideal* or the *effects* focused simulations, and like in Bramley, Lagnado, and Speekenbrink (2015), learners favoured "one-on" tests. However, in line with an *edge* or *confirmation* they also selected a substantial number of "two-on" and "one-on, one-off" interventions, doing so on early as well as late tests while the ideal learner only predicted using "one-on, one-off" tests on the last few trials. As in Bramley, Lagnado, and Speekenbrink (2015) and consistent with *confirmatory* testing, participants were more likely to fix "on" components with at least one child according to their latest hypothesis b^{t-1}: 49% compared to 30% $t(238) = 5.5, p < .001$. The overall pattern was not clearly consistent with any one local focus but might be consistent with a mixture of all three.

5.7 Experiment 5: Unknown strengths

In this experiment, we focused on cases where participants are not pretrained on $\mathbf{w}$ (see Appendix A for the computational level details of how to incorporate uncertainty over $\mathbf{w}$ in model inference and intervention choice).

We took advantage of the fact that participants would experience substantially greater uncertainty given ignorance about $\mathbf{w}$ to assess their ability to estimate local uncertainty based on recently observed data $\mathcal{D}_r^t$ in order to choose where to focus subsequent tests. This is central to any scheme for intervention selection. Thus, in Experiment 3, we elicited the participants' confidence about the edges in each judgment. If participants track local uncertainties based on recent evidence, we should expect these to correlate with uncertainties given $\mathcal{D}_r^t$. In particular, given the representation associated with *Neurath's ship*, we might also expect the local confidences to be evaluated while leaning on

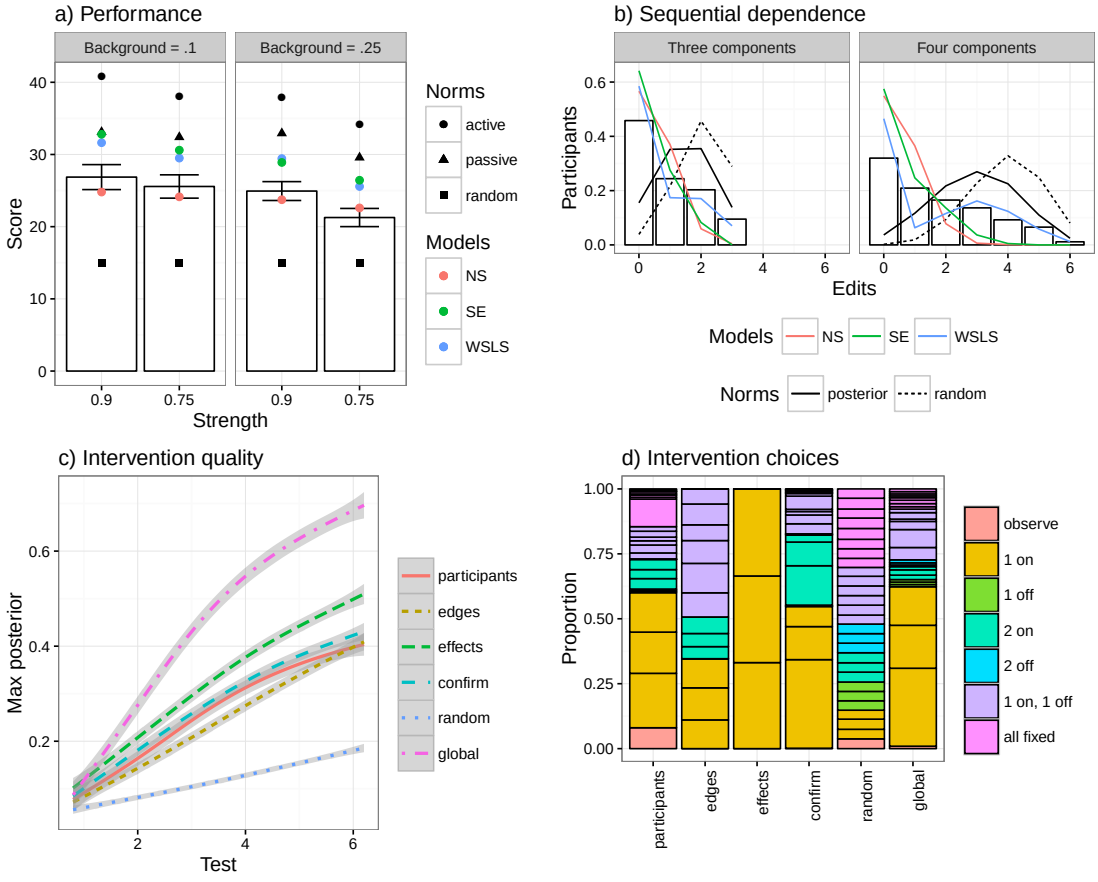


FIGURE 5.6: Experiment 4: Learning larger causal models; performance and interventions. a) Accuracy by condition. Bars are participants and points compare with the models. b) Sequential dependence as in Figure 5.3b. c) Quality of participants' and simulated learners' intervention choices in the three variable problems as in Figure 5.3c. d) The proportion of interventions of different types chosen by participants compared to simulated learners in the three variable problems, otherwise as in Figure 5.3d.

the rest of the model for support. This means they should reflect conditional uncertainty in the edge $H(E_{ij}|E_{\setminus ij}, \mathcal{D}_r^t; \mathcal{C}_r^t)$ more closely than the marginal uncertainty $H(E_{ij}|\mathcal{D}_r^t; \mathcal{C}_r^t)$ which involves averaging across all the possible states of the other edges.

We also elicited predictions about the outcome of each chosen test before the outcome was revealed. If participants maintained only a single hypothesis, we expected this to be reflected in their predictions. Thus, for a *Neurath's ship* learner, it would be predominantly the predictive distribution under their current hypothesis rather than the average across models.

Finally, in Bramley, Lagnado, and Speekenbrink (2015) and Experiment 4, participants' intervention selections showed hints of being motivated by a mixture of local aspects of the overall uncertainty, with overall patterns most consistent with focus on a mixture of different local aspects of uncertainty. To test this idea more thoroughly, in the final

problem in Experiment 5 we explicitly probed participants' beliefs about their intervention choices through eliciting free responses which we go on to code and compare to our model predictions.

5.7.1 Methods

Participants

111 UCL undergraduates (mean ± SD age 18.7 ± 0.9, 22 male) took part in Experiment 5 as part of a course. They were incentivised to be accurate based on randomly selected trials as before, but this time with the opportunity to win *Amazon*™ vouchers rather than money. Participants were split randomly into 8 groups of mean size 13.8 ± 3.4, each of which was presented with a different condition in terms of the value of $\mathbf{w}$ and the way that they had to register their responses.

Design and procedure

Experiment 5 used the same task interface as the other experiments, but focused just on the three variable problems (devices 1-5) and an additional device (6) in which none of the components was connected (Figure 5.8). Like in Experiment 4, there were two causal strength conditions $w_S \in [0.9, 0.75]$ and two background noise conditions $w_B \in [0.1, 0.25]$. However, unlike Experiment 4, participants were not trained on these parameters, but only told that: *"the connections do not always work"*, and *"sometimes components can activate by chance"*.

To assess the extent the different reporting conditions drove lower sequential dependence in Experiment 4 relative to Bramley, Lagnado, and Speekenbrink (2015), we examined two reporting conditions between subjects: *remain* and *disappear*. In the *remain* condition, judgments stayed on the screen into the next test, so participants did not have to change anything if they wanted to register the same judgment at t as at $t-1$. In the *disappear* condition, the previous judgment disappeared as soon as participants entered a new test. They then had to explicitly make a choice for every connection after each test.

In addition to the structure judgments and interventions, we also elicited additional probability measures from participants. First, after selecting a test, but before seeing the outcome, participants were asked to predict what would happen to the variables

they had left free. To do this they would set a slider for each variable they had left free to vary. The left pole of the slider was labelled "Sure off", the right pole "Sure on"and the middle setting indicated maximal uncertainty (Figure 5.7a). Second, after drawing their best guess about the causal model by setting each edge between the variables, participants were asked how sure they were about each edge. Again they would respond by setting a slider, this time between "Guess" on the left indicating maximal uncertainty and "Sure" on the right indicating high confidence that that edge judgment was correct (Figure 5.7b). Participants were trained and tested on interpretation of the slider extremes and midpoints in an additional interactive page during the instructions.

Participants faced the six devices in random order, with six tests per device followed by feedback as in Experiment 4. Then they faced one additional test problem. On this problem, the true structure was always a Chain (Figure 5.8, device 7). On this final problem, participants did not have to set sliders. Instead, after they selected each test, but before seeing its outcome, they were asked why they had selected that intervention. Labels would appear on the nodes and participants were invited to *"Explain why you chose this combination of fixed and unfixed components. Use labels 'A' 'B', 'C' to talk about particular components or connections"* in a text box that would appear below the device. Responses were constrained to be at least 5 characters long. The Chain (device 3) was chosen for this problem because in Bramley, Lagnado, and Speekenbrink (2015) and Experiment 4, participants often did not select the crucial $\mathrm{Do}[X_1 = 1, X_2 = 0]$ intervention that would allow them to distinguish a Chain from a Fully-connected model (device 5) making this an interesting case for exploring divergence between participants' behaviour and ideal active learning.

Finally, at the end of the experiment participants were asked to estimate the reliability w_S of the true connections: *"In your opinion, how reliable were the devices? i.e. How frequently would fixing a cause component ON make the effect component turn ON too?"* and the level of background noise w_B: *"In your opinion, how frequently did components activate by themselves (when they were not fixed by you, or caused by any of the device's other components)?"* by setting sliders between *"0% (never)"* and *"100% (always)"*.

A demo of Experiment 5 can be viewed at `http://www.ucl.ac.uk/lagnado-lab/el/nbt`.

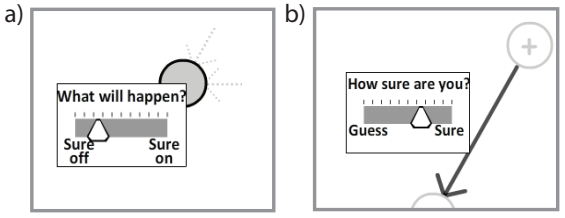


FIGURE 5.7: Experiment 5: Unknown strengths; additional measures - a) Outcome expectation sliders b) Edge confidence sliders.

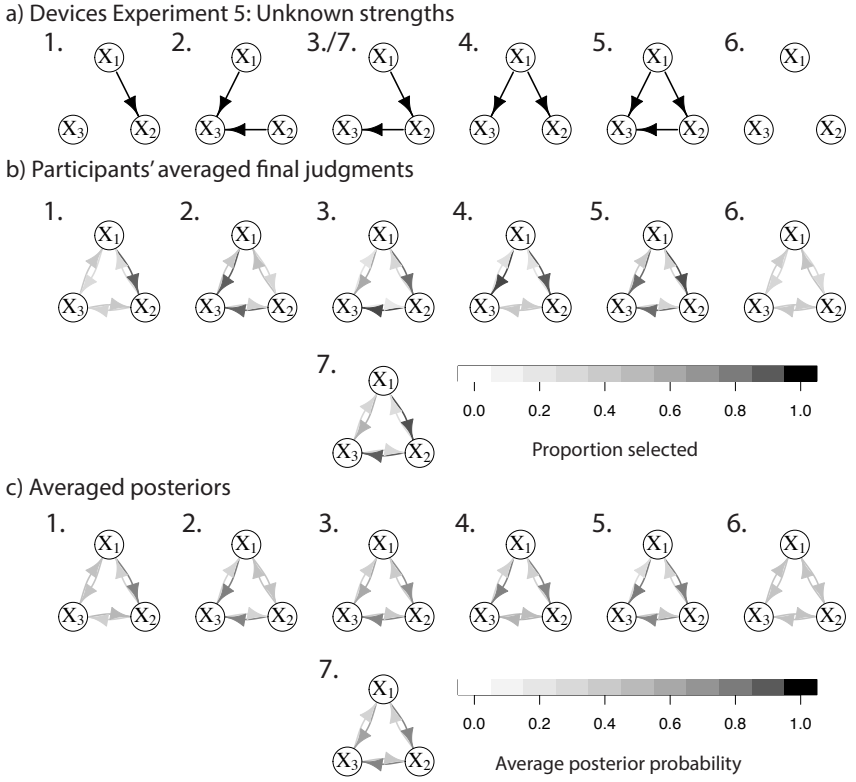


FIGURE 5.8: Experiment 5: Unknown strengths; true models and final judgments. a) The true models faced by participants. b) Weighted average final judgments by participants. Darker arrows indicate that a larger proportion of participants marked this link in their final model. Note that problem 7 was a repeat of the Chain (problem 3) with the write aloud protocol. c) Bayes-optimal marginal probability of each edge in $\int_{\mathbf{w}} P(M|D^T; C^T)p(\mathbf{w})\, \mathrm{d}\mathbf{w}$ averaged over participants' data assuming a uniform independent prior over $\mathbf{w}$.

5.7.2 Results and discussion

Judgments

As in the experiments where participants were trained on $\mathbf{w}$, accuracy was significantly higher than chance in all conditions (all 8 t statistics > 6.1 all p values < 0.001) and lower than a Bayes optimal observer observing the same data as them. Because the

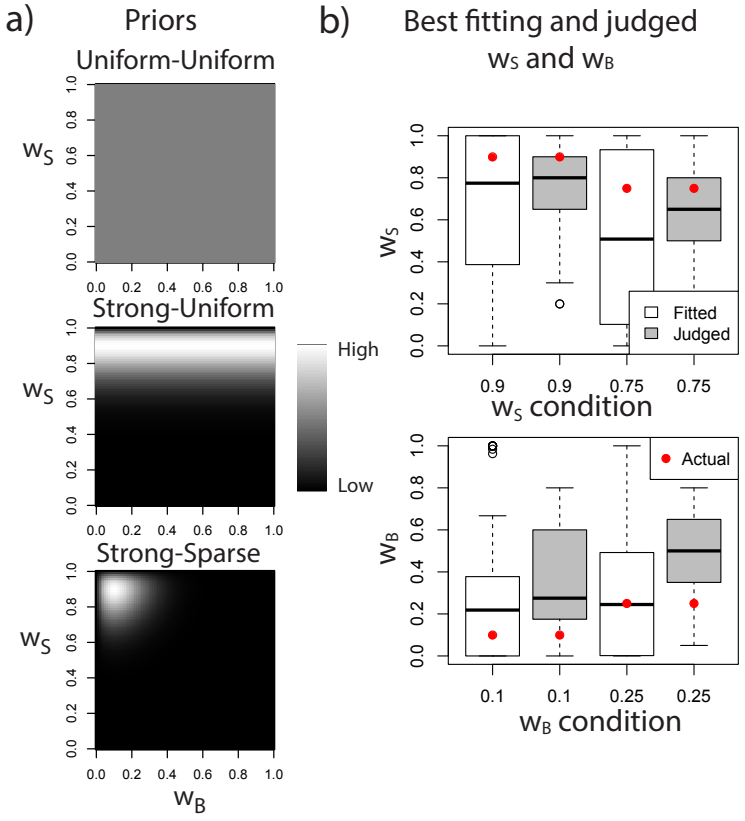


FIGURE 5.9: Experiment 5: Unknown strengths; priors and strength estimates. a) Visualisations of Uniform-Uniform, Strong-Uniform and Strong-Sparse priors on w_S and w_B b) Participants final judgments of amount of background noise (w_B) and strength (w_S), rescaled from 100 point scale to 0-1, and best-fitting w_S^* and w_B^* estimates assuming ideal Bayesian updating. Boxplot shows the medians and quartiles.

noise was unspecified, we explored several reasonable priors on $\mathbf{w}$ (always assuming that w_S and w_B were independent) when computing posteriors. Firstly, we considered a *uniform-uniform* prior that made no assumptions about either w_S or w_B (UU) where $\mathbf{w} \sim \text{Uniform}(0,1)^2$. We also considered a *strong-uniform* (SU) variant, following Yeung and Griffiths (2011), expecting causes to be reliable — $w_S \sim \text{Beta}(2,10)$, but making no assumptions about background noise — $w_B \sim \text{Uniform}(0,1)$. Additionally, we considered a *sparse-strong* (SS) variant following Lu et al (2008), encoding an expectation of high edge reliability — $w_S \sim \text{Beta}(2,10)$, and relatively little background noise — $w_B \sim \text{Beta}(10,2)$.[10] The choice of parameter prior made little difference to the Bayes optimal observer's judgment accuracy. Thus, participants significantly underperformed the Bayes optimal observer in all conditions regardless of the assumed prior, except for condition $w_S = 0.75; w_B = 0.1$, *remain*) under the SU prior, and $w_S = 0.75; w_B = 0.25$, *remain* under all three considered priors.

[10] Lu et al. (2008) actually used a joint prior $p(w_S, w_B)$ with density concentrated in the top left and bottom right, as in edges *either* had high strength and low background noise or high background noise and low strength.

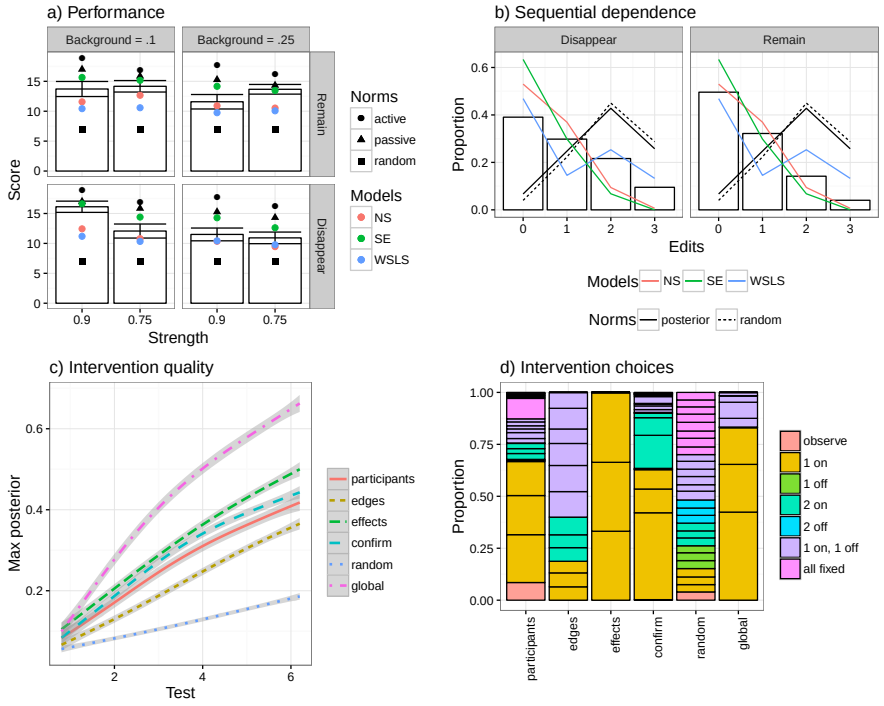


FIGURE 5.10: Experiment 5: Unknown strengths; performance and interventions. Subplots as in Figure 5.6.

Comparison with known strength experiments

Performance in Experiment 5 was comparable to the 3-variable problems in Experiment 4 where the underlying **w** conditions were identical. Mean accuracy was actually slightly higher 0.61 ± 0.21 compared to 0.56 ± 0.21 for the matched problems in Experiment 4, although not significantly so $t(229) = 1.9, p = .054$. This suggests that participants were able to make reasonable structure judgments without knowledge of the exact parameters. We found that participants' final judgments of w_S and w_B and best fitting estimates assuming rational updating w_S^* and w_B^* suffered bias and variance (Figure 5.9 b).[11]

As with Experiment 4 and 5, participants were not affected by the reliability of the connections themselves w_S $t(106) = 0.88, p = 0.37$, but were affected by higher levels of background noise w_B $t(108) = 2.7, p = 0.008$. There was no difference in performance between the two judgment elicitation conditions $t(108) = 0.67, p = 0.50$.

Participants were no more or less accurate on the final problem when identifying a Chain structure for the second time (device 7). The most frequent error once again was mistaking the Chain structure for the Fully-connected structure, made by 17/111

[11] Fifty-eight participants' final w_B judgments were incorrectly stored, so the N for w_B judgments was 53 rather than 111.

participants, although this was reduced to 11/111 when facing the Chain structure again on device 7, with only a single participant making the same error twice.

Average edit distance between sequential judgments about the same device was significantly increased by removing the record of previous judgments between trials, going from .73 in the *remain* condition to 1.0 in the *disappear* condition $t(109) = 3.5, p < .001$. Edit distances even in the *disappear* condition were still significantly lower than those predicted by UU, SU or SS posterior or random sampling (all p's $< .001$). As in Experiment 4 there was a strong negative relationship between number of edits and performance $F(1, 109) = 102, \beta = 6.4, \eta^2 = .48, p < .001$. The edit-distance–performance relationship interacted weakly with condition $t(108) = 1.9, \beta = 1.3, p = .049$ becoming stronger in the *disappear* condition. Again, the 71 participants who scored significantly above chance ($\frac{12}{21}$ or higher by χ^2 test) had lower edit distances of 0.66 ± 0.29 than the remaining 40 participants' 1.3 ± 0.44.

Additional measures

Participants' edge confidence judgments increased significantly over trials $\chi^2(1) = 2060, \beta = .04, SE = .0008, p < .001$, going from $.57 \pm .20$ on the first trial to $.78 \pm .19$ by the final trial.The probability of changing an edge at the next time point was weakly inversely related to the learners' reported confidence in it $\chi^2(1) = 67, \beta = -.03, SE = .004, p < .001$. Reported edge confidences were correlated with both the conditional probability of the edge states given the rest of the current model $r^{cond} = .20$ and the *marginal* probability of the edge-state in the full posterior under the UU prior $r^{mar} = .17$ but these correlations did not differ significantly.

As predicted, reported outcome predictions were more closely related to the predictive distribution under the participants' latest structure judgment b^{t-1}: $\chi^2(1) = 1044, \beta = .35, SE = .010, p < .001$ than marginalised over the full posterior $\chi^2(1) = 580, \beta = .29, SE = .012, p < .001$. The latest-structure to prediction relationship was significantly stronger than the marginal posterior to prediction relationship by Cox test $Z = 10.9, p < .001$.

Interventions

The overall distribution of intervention choices was broadly similar to the other Experiments (Figure 5.10). "One-on" interventions were the most frequently chosen, making up

39% of selections. However, unlike the previous Experiments, and consistent with *edge* focused learning, the constrained "one-on one-off" interventions were almost as common as single "one-on" interventions, making up 38% of tests compared to 12% across 3-variable problems in Experiment 4. The intervention selections and informativeness of intervention sequences were not closely consistent with global expected information, nor any single type of local focus, but could again be consistent with a mixture of local *effect* focused, *edge* focused and *confirmation* focused queries.

Free explanations

For device 7, participants gave free explanations for their intervention choices on each of their six tests. The overall distribution of intervention choices did not differ significantly from the original presentation of the Chain (device 3) $\chi^2 = 31, p = 0.21$ suggesting that the different response format did not affect the intervention choices that participants made. In order to assess what the explanations tell us about participants' intervention choices, we asked two independent coders to categorise the free responses into 8 categories. The categories were chosen in a partly data-driven, partly hypothesis-driven way: 1. An initial set of categories were selected, with the goal of distinguishing the approximations introduced in *A local uncertainty schema* from global strategies like uncertainty sampling or expected information maximisation. 2. A subset of the data was then checked and the categories were refined to better delineate their responses with minimal membership ambiguity.

The eight resulting categories were:

1. The participant just wanted to learn about one specific connection. [Corresponding to *edge focused* testing]
2. The participant wanted to learn about two specific connections.
3. The participant wanted to learn about all three connections. [Corresponding to *globally focused* testing]
4. The participant wanted to learn what a particular component can affect but did not mention a specific pattern of connections. [Corresponding to *effect focused* testing]
5. The participant wanted to test / check / confirm their current hypothesis. [Corresponding to *confirmatory* testing]

6. The participant wanted to learn about the randomness in the system (as opposed to the location of the connections). [Corresponding to a focus on learning about noise rather than structure]

7. The participant chose randomly / by mistake / to use up unwanted tests / they say they did not understand what they are doing /it is clear they were not engaging with the task.

8. The participant's explanation was complex / underspecified / did not seem to fall in any of the above categories.

A supplementary file (available at `http://www.ucl.ac.uk/lagnado-lab/el/nbt`) contains all the materials given to coders and the full set of participant responses. Coders were permitted to assign more than one category per response, but had to select a primary category. When the category referred to particular component label(s), the rater would record these, and when it referred to a specific connection they would record which direction (if specified) and the components involved. These details will be used to facilitate a quantitative comparison between participants' explanations and our model fits in the next Section. Raters normally just selected one category per response, selecting additional categories on only 8% of trials. Inter-rater agreement on the primary category was 0.73, and Cohen's $\kappa = 0.64 \pm 0.04$, both higher than their respective heuristic criteria for adequacy of 0.7 and 0.6 (Krippendorff, 2012; Landis & Koch, 1977).

Figure 5.11, shows the proportion of responses in the different categories across the six trials. On the first trial participants were most likely to be categorised as 4. — focused on identifying what a particular variable could effect. On subsequent trials they most frequently categorised as 1. — focusing on learning about a specific connection. Toward the end, explanations became more diverse and were increasingly categorised as 5. — confirmatory testing or 6. learning about the noise in the system. Individuals almost always gave a range of different explanations across their six tests, falling under 3.0 ± 0.99 different categories on average, with only 5/111 participants providing explanations from the same category all six times (3 all-fours, 1 all-threes. and 1 all-eights).

Explanation type was predictive of performance $F(8, 657) = 13.75, \eta^2 = 0.14, p < 0.001$. Taking category 7 — unprincipled or random intervening — as the reference category with low average performance of 10.2 points out of a possible 21, categories 1,2,4,5, and 6 were all associated with significantly higher final scores $[14.5, 12.9, 13.9, 13.9, 13.9]$ points, all p's < 0.001.

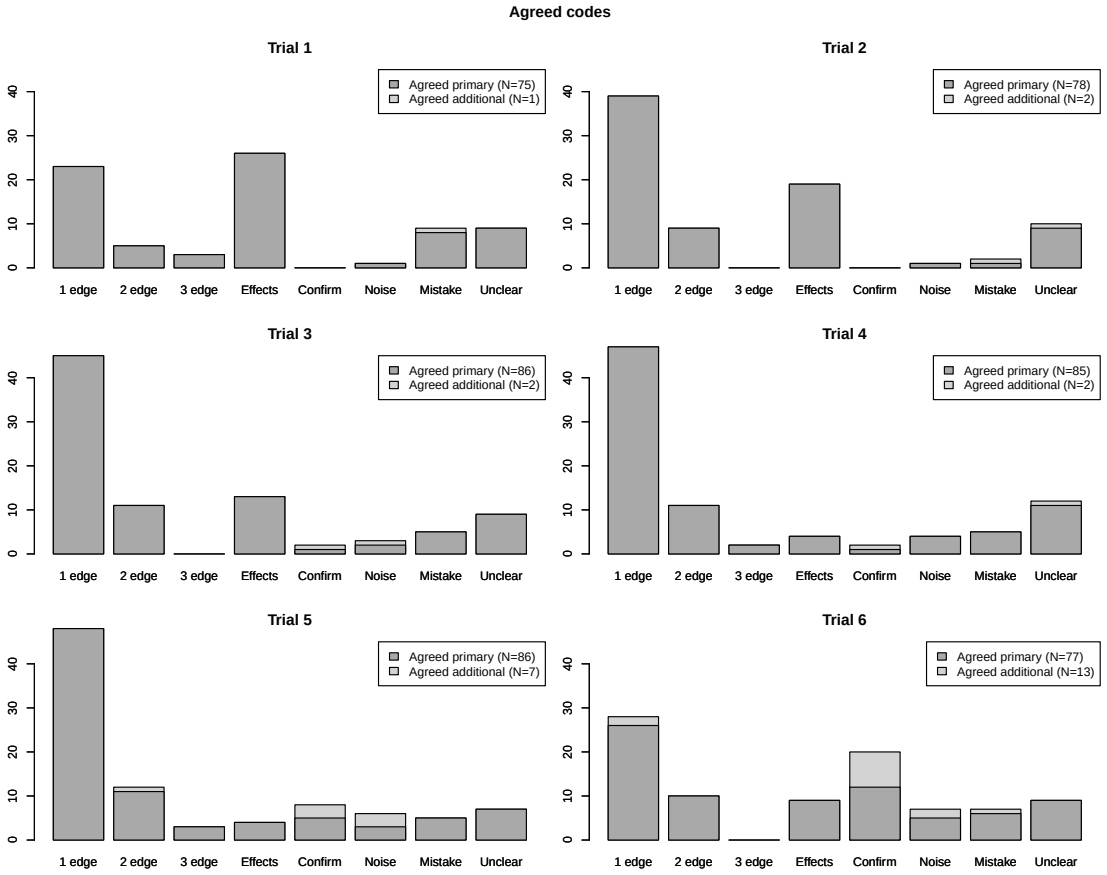


FIGURE 5.11: Experiment 5: Unknown strengths; free explanations for interventions agreed codes over the six tests in problem 7.

5.7.3 Summary of Experiments

In all these experiments, participants were clearly able to generate plausible causal models but also did so suboptimally. Averaged across participants, final model judgments resembled the posterior over models (e.g. Figures 5.5c and 5.8c), however individuals' trajectories typically exhibited strong sequential dependence, with the probability of moving to a new model decreasing with its edit distance from the previous model. This is consistent with our hypothesis that individuals normally maintain a single hypothesis and update it piece by piece. As found in previous research, participants were worst at separating the direct and indirect causes in the Chain (3; 8) and Fully-connected (5; 10) models. A closer look at participants' intervention choices suggests that this was due to a common failure to generate the constrained interventions, such as $\mathrm{Do}[X_1 = 1, X_2 = 0]$, necessary to disambiguate these options. The *simple endorser* model predicts this error by proposing that people ignore the dependencies between the different edges. Our framework provides a more nuanced explanation. Whether a learner will correctly disambiguate these options depends on whether they focus on $X_1 - X_3$ before or after

having inferred $X_1 \to X_2$ and $X_2 \to X_3$. If they consider $X_1 - X_3$ after, then they will tend to fix X_2 "off", realising it is necessary to prevent the indirect path from confounding the outcome of their test. However, if they have no connection marked from X_1 to X_2 or from X_2 to X_3, they will not expect this confounding activation and so have no motivation to fix X_2 "off" when testing $X_1 - X_3$.

Participants' overall distributions of intervention selections resembled a mixture of *edge*, *effect* and *confirmation* focused testing, but their distributions of choices were relatively invariant across conditions and trials while the efficient learners' were much more dynamic. Comparison with the final global information gathered revealed that they did not select which variables to target particularly efficiently, leading to a considerable discrepancy between the total information gathered by participants compared to an ideal active learner. However, participants also displayed hints of adaptation of strategy over the trials: with a preference for confirmatory testing, being more likely to fix variables "on" when they had children according to their latest hypothesis b^{t-1}, and displaying a modest shift toward more constrained interventions in later trials.

In Experiment 5 we saw that people were able to identify causal structure effectively without specific parameter knowledge. Comparing a range of plausible prior assumptions about edge reliability w_S and the level of background noise w_B yielded little difference in judgment or intervention choice predictions. Participants' overall judgment accuracy was not affected by the *remain/disappear* reporting condition, but this did affect sequential dependence, especially for lower performers who may have often forgotten their previous judgment when making their next one. The idea, common to the three judgment rules we consider, that people represent one model at a time was also supported by the additional measures elicited from participants during the task. With a single hypothesis rather than distributional beliefs, intervention outcome predictions could only be generated by the current hypothesis rather than averaged and weighted over all possible models. Consistent with this idea, we found participants' expectation judgments were more in line with their current hypothesis than the marginal likelihoods, although we note that these measures were quite noisy and the effects quite small.

5.8 Modelling individual behaviour

Across all three examined experiments we found a qualitative correspondence, both between our *Neurath's ship* simulations and participants' judgments, and between the two

stage local intervention schema and participants' interventions. However, both *simple endorsement* and *win-stay, lose-sample* also appeared to do a good job of capturing qualitative judgment patterns. In order to validate quantitatively which of these models better describes participants' behaviour, we fit the models to the data and assessed their competence relative also to *win-stay, lose-sample* and *simple endorsement*. By fitting the models separately to individual participants we also assessed individual differences in learning behaviour, and thus gained a finer-grained picture of the processes involved.

5.8.1 Judgments

Models

We compared six models to participants' judgments, the three process models we considered in the experiment *Neurath's ship* (NS), *simple endorser* (SE), *win-stay,lose-sample* (WSLS), alongside an efficient Bayesian learner (*Rational*) and two null models *Baseline* and NS-RE.

For NS, we fit three parameters:

1. An average search length parameter λ controlling the probability of searching for different lengths k on each belief update.
2. A search behaviour parameter ω controlling how strongly the learner moves toward the more likely state for an edge when updating it (recalling that $\omega = 1$ leads to probability matching, while $\omega = \infty$ leads to deterministic hill climbing and $\omega = 0$ to making random local edits).
3. A lapse parameter ϵ controlling a mixture between the model predictions and a uniform distribution.

Including the last parameter into equation 5.1, this resulted in the following equation

$$P(b^t = m|\mathcal{D}_r^t, \mathcal{C}_r^t, \mathbf{w}, b^{t-1}, \omega, \lambda) = (1-\epsilon)\sum_0^\infty \frac{\lambda^k e^{-\lambda}}{k!}[(R_t^\omega)^k]_{b^{t-1}m} + \frac{\epsilon}{|M|} \tag{5.9}$$

where R is a Markov matrix expressing the options for local improvement.

We operationalised the *Simple endorser* (SE) (Bramley, Lagnado, & Speekenbrink, 2015) with two parameters. One is the probability $\rho \in [0, 1]$ with which the belief state is updated from b^{t-1} to include extra edges from any currently fixed "on" node(s) to any

activated nodes and to exclude edges from any currently fixed "on" node(s) to any non-activated nodes (we write b^{t-1}_{+SE}). With the complementary probability $1-\rho$, it stays the same as b^{t-1}. As with the NS model we also included a lapse parameter mixing in a probability of choosing something at random, giving

$$P(b^t = m|\mathbf{d}, \mathbf{w}) = (1-\epsilon)(\rho\ b^{t-1}_{+SE} + (1-\rho)b^{t-1}) + \frac{\epsilon}{|\mathcal{M}|} \tag{5.10}$$

Win-stay, lose-sample (WSLS) (Bonawitz et al., 2014) predicts that participants stick with their current model b^{t-1} with probability $p(\mathbf{d}^t|b^{t-1}, \mathbf{w}, c^t)$ or else draw a sample from the full posterior with probability $1 - p(\mathbf{d}^t|b^{t-1}, \mathbf{w}, c^t)$. The fitted version of this model had a single lapse parameter ϵ giving

$$P(b^t = m|D^t, \mathbf{w}) = (1-\epsilon)\Big((1-P(\mathbf{d}^t|b^{t-1}, \mathbf{w}, c^t))P(M|D^t, \mathbf{w})_t + P(\mathbf{d}^t|b^{t-1}, \mathbf{w}, c^t)[m = b^{t-1}]\Big) + \frac{\epsilon}{|\mathcal{M}|} \tag{5.11}$$

The final model, *Rational* was a variant of the Bayes-optimal observer (Section 2) that attempted to select the maximum a posteriori causal structure $\max P(M|D^t, \mathbf{w}; C^t)$ with each judgment, with a soft maximisation (Luce, 1959) governed by inverse temperature parameter θ and a lapse parameter ϵ. For this, we considered

$$P(b^t = m|D^t, \mathbf{w}) = (1-\epsilon)\frac{\exp(P(M|D^t, \mathbf{w})_t\theta)}{\sum_{m' \in M}\exp(P(m'|D^t, \mathbf{w})_t\theta)} + \frac{\epsilon}{|\mathcal{M}|} \tag{5.12}$$

Baseline is a parameter-free baseline that assumes each judgment to be a random draw from all possible causal models

$$p(b^t = m) = \frac{\epsilon}{|\mathcal{M}|} \tag{5.13}$$

(leading to a probability of approximately $\frac{1}{3}$ for each edge).

One concern with this baseline is that judgments might exhibit sequential dependence yet be unrelated to data $\mathcal{D}^t_r$. Therefore we also considered a baseline variant of the NS model in which the search behaviour parameter ω was fixed to 0, resulting in a (R)andom (E)dit model (NS-RE) that walks randomly around the hypothesis space for k steps on each update. For this model, small k simply denotes more inertia.

Each of these belief models output a likelihood based on the probability the model assigns to a belief of b^t, given the most recent outcome $\mathbf{d}^t$ (SE), outcomes since the last belief change $\mathcal{D}_r^t$ (NS), or all outcomes D^t (WSLS, *Rational*), and the most recent judgment b^{t-1}. Because the choice of prior for Experiment 5 made negligible difference to our results, we only report models assuming uniform (UU) priors on $\mathbf{w}$. For Experiment 5, we also marginalised over the unknown values of $\mathbf{w}$ rather than conditioning as in the other experiments as detailed in Appendix B.

Evaluation

To compare these models quantitatively, we used maximum likelihood optimisation as implemented by R's `optim` function to fit the model separately to each of the 370 participants across all three experiments.[12] We used the Bayesian Information Criterion (BIC, Schwarz, 1978) to compare the models while accommodating their differing numbers of parameters. *Baseline* acts as the null model for computing BICs and pseudo-R^2s (Dobson, 2010) for the other models. In Bramley, Lagnado, and Speekenbrink (2015) participants were not forced to select something for each edge immediately, although once they did so they could not return to "unspecified", and they could also respond with cyclic causal model if they wanted. Therefore, we fit only the 75% of tests where the participants report a fully specified non-cyclic belief, taking the b^{t-1} to be the unconnected model on the first fully specified judgment, as we do with b^0 in the other experiments. Recalculating the transition probabilities on the fly in the optimisation of ω was infeasibly computationally intensive for the four-variable problems. So for Experiment 4 we first fit all three parameters to the three-variable problems only, then used the best fitting ω parameters from this fit when fitting the λ and ϵ on the full data. In Bramley, Lagnado, and Speekenbrink (2015) and Experiment 5 we were able to fit all three parameters.

Results and discussion

Table 5.3 details the results of the model fits to all experiments. Summed across all participants, NS has the lowest total BIC (93381) with the SE in second place (94326), followed by WSLS with (97643), then NS-RE (101837), *Rational* (1207209) and finally *Baseline* with (149313). NS was also the best fitting model for Bramley, Lagnado, and Speekenbrink (2015) and Experiment 4, with SE winning in Experiment 5. Thus,

[12] In Appendix B we provide additional detail on how the models were fit.

Table 5.3: Belief Model Fits.

Exps 1–2 (Bramley, Lagnado, & Speekenbrink, 2015)															
Model	λ M	λ SD	ω M	ω SD	θ M	θ SD	ρ M	ρ SD	ϵ M	ϵ SD	N fit	M acc	logL	R^2	BIC
Baseline											1	0.27	-17836	0	35672
NS-RE	0.17	51.5							0.21	0.19	32	0.36	-9379	0.49	19762
SE							0.18	0.18	0.21	0.24	13	0.60	-8819	0.53	18642
WSLS									0.14	0.28	56	0.85	-9117	0.52	18736
NS	**1.20**	**100.6**		**243.8**					**0.05**	**0.20**	**27**	**0.66**	**-8197**	**0.56**	**17901**
Rational					5	124.8			0.00	0.39	10	0.93	-12089	0.36	25182
Exp 4: Learning larger models															
Model	λ M	λ SD	ω M	ω SD	θ M	θ SD	ρ M	ρ SD	ϵ M	ϵ SD	N fit	M acc	logL	R^2	BIC
Baseline											1	0.33	-41814	0	83628
NS-RE	1.23	1.5							0.25	0.28	33	0.55	-29235	0.30	59490
SE							0.51	0.32	0.56	0.29	30	0.49	-27736	0.34	56492
WSLS									0.32	0.34	29	0.52	-28772	0.31	58053
NS	**1.63**	**1.7**	**4**	**241.0**					**0.20**	**0.29**	**27**	**0.56**	**-27234**	**0.35**	**55896**
Rational					13	143.2			0.14	0.36	0		-36362	0.13	73743
Exp 5: Unknown strengths															
Model	λ M	λ SD	ω M	ω SD	θ M	θ SD	ρ M	ρ SD	ϵ M	ϵ SD	N fit rem/diss	M acc	logL	R^2	BIC rem/dis
Baseline											0/2	0.21	-15006	0	14330/15682
NS-RE	0.9	3							0.10	0.29	3/2	0.36	-10877	0.28	10050/12535
SE							**0.80**	**0.30**	**0.43**	**0.27**	**21/16**	**0.75**	**-9181**	**0.39**	**8257 /10936**
WSLS										0.31	9/8	0.50	-10220	0.32	9402/11452
NS	1.8	121	9.2	247					0.03	0.20	13/18	0.55	-9170	0.39	8620/10964
Rational					31.5	282				0.26	7/12	0.70	-10482	0.30	10116/11678

Note: Columns: M = median estimated parameters across all participants, SD = standard deviation of parameter estimate across all participants, N fit = number of participants best fit by each model, M acc = average proportion of edges identified correctly by participants best fit by this model, LogL = total log likelihood of model over all participants, R^2 = median McFadden's pseudo-R^2 across all participants, BIC = aggregate Bayesian information criterion across all participants. For Exp 5, rem = remain condition, dis =disappear condition. Best fitting model denoted with boldface.

all three heuristics substantially beat an exact Bayesian inference account of causal judgment here, but *Neurath's ship*, with its ability to capture a graded dependence on prior beliefs, outperformed WSLS substantially, and the heuristic SE to a lesser degree. In terms of number of individuals best fit, Table 5.3 shows a broad spread across models: WSLS – 102, NS – 85, SE – 80, NS-RE – 70, *Rational* – 28 , *Baseline* – 4.

The diversity of individual fits across strategies raises the question of the identifiability of the different models. To assess how reliably genuine followers of the different proposed strategies would be identified by our modelling procedure, we simulated participants using the fitted parameters for each model for each of the actual participants in all three examined experiments. We then fit all six models to these simulated participants and report the rates at which simulations are best-captured by each model. Table A.1 in the Appendix provides the complete results for this recovery analysis. Overall, the generating model was recovered 74% of the time for Bramley, Lagnado, and Speekenbrink (2015), 82% for Experiment 4 and 75% for Experiment 5 (chance would be

17%). In all three experiments, data generated by *Baseline*, WSLS and SE was nearly always correctly recaptured, indicating that we can treat cases where participants are well described by these models as genuine. Additionally NS almost never captured data generated by any of the other models, providing reassurance that NS is not simply fitting participants who are doing something more in line with SE or WSLS. However, data actually generated by NS was frequently recaptured by the NS-RE (random edit) null model that makes NS-style local edits but does not preferentially approach more likely models. This was true in the majority of cases in Bramley, Lagnado, and Speekenbrink (2015) and Experiment 5. Some of the cases where NS-RE captures NS-generated simulations are based on participants who were better described by NS-RE in the first place (e.g. whose search behaviour was too random to justify ω's inclusion). We find a similar effect whereby simulated *Rational* participants with relatively low θs or high ϵs are more parsimoniously described by *Baseline*. This is supported by looking at the more complex four variable problems in Experiment 4: NS simulations were identified the majority of the time, and when restricted to simulations based on parameters from participants who were actually best described by NS, 24/27 were recovered successfully. Thus, it is plausible that some of the 70 NR-RE participants were in fact doing something more in line with NS. There is a suggestion of this in Experiment 4, where the mean accuracy of the NS-RE participants is commensurate with SE, WSLS and NS.

The performance of a handful of participants — 10 in Bramley, Lagnado, and Speekenbrink (2015), and 19 in Experiment 5 — were best fit by the *Rational* model, which has one fewer parameter than NS. Naturally, these participants performed particularly well, scoring near ceiling in Bramley, Lagnado, and Speekenbrink (2015) (identifying 14.0 of the 15 connections) and as high as the ideal learning simulations in Experiment 5 — 14.7/21 compared to an average of 15.5 for perfect Bayesian integration. This, along with the lower recovery rates for these experiments, suggests that their design — both being motivated primarily to look closely at intervention choice — may not have been difficult enough to separate the process from the normative predictions about the judgments.

Figures 5.12a and b show the range of the fitted λ and ω parameters under NS. In line with our predictions, participants' average fitted search lengths (λ) were mainly small, with medians between 1 and 2 in all three experiments.[13] Because this parameter merely encodes a participant's average search length this means that the same participant would

[13] A few participants made judgments that were sequentially anti-correlated leading to λ parameters at the limit of the optimisation routine's precision and correspondingly large standard deviations in Experiment 4 and 5.

sometimes not search at all, staying exactly where they are ($k = 0$), or might also sometimes search much longer (e.g. $k \gg \lambda$). The median fitted ωs of 6, 4 and 9.2 across the three experiments are suggestive of moderate hill-climbing. A substantial number of participants had very large values of ω indicative of near-deterministic hill climbing. We discuss this trade-off further in the General Discussion. However, note that we were only able to fit these values to the easier three variable problems. It might be that the largest values would have been tempered if they could have been fit to the four variable problems as well.

5.8.2 Interventions

Models

We compared our local model of intervention choice (Section 4) to a globally-focused and a baseline model. Each intervention model output a likelihood for an intervention choice of $\mathbf{c}^t$, depending on $\mathcal{D}_r^t, \mathcal{C}_r^t$ and b^{t-1}.

We compared the overall distribution of participants' intervention selections and final performance with *edge* focused, *effect* focused and *confirmation* focused tests. We found that none of these models alone closely resembled participants' response patterns, but overall distributions were consistent with a mixture of different types of local tests. This was also supported by the free-response coding in Experiment 5, showing that participants would typically report targeting a mixture of specific edges, effects of specific variables and confirming the current hypothesis. Therefore, we considered four locally driven intervention selection models, one for each of the three foci, plus a mixture.

For the *edge* model, the possible foci $\mathcal{L}$ included the 3 (or 6) edges in the model. For the *effect* model, it comprised the 3 components (or 4 in the 4-variable case). The confirmation model always had the same focus — comparing b^t to null b^0 of no connectivity. The *mixed* model contained all 7 (or 11) foci. As in Equations 5.4 and 5.5 in Section 4, each model would first compute a softmax probability of choosing each possible focus $\mathbf{l}^t \in \mathcal{L}$. Within each chosen focus it would also calculate the softmax probability of selecting each intervention, governed by another inverse temperature parameter $\eta \in [0, \infty]$. The total likelihood of the next intervention choice was thus a soft-maximisation-weighted average of choice probabilities across possible focuses

$$P(\mathbf{c}^t|\eta, \rho, \mathcal{D}_r^t, b^{t-1}, \mathbf{w}) = \sum_{\mathbf{l}\in L} P(\mathbf{c}|\mathbf{l}, \eta, b^{t-1}, \mathbf{w}) \frac{\exp(H(\mathbf{l}|\mathcal{D}_r^t, b^{t-1}, \mathbf{w}; \mathcal{C}_r^t)\rho)}{\sum_{\mathbf{l}'\in\mathcal{L}} \exp(H(\mathbf{l}'|\mathcal{D}_r^t, b^{t-1}, \mathbf{w}; \mathcal{C}_r^t)\rho)} \quad (5.14)$$

where

$$P(\mathbf{c}|\mathbf{l},\eta,b^{t-1},\mathbf{w}) = \frac{\exp\left(\mathbb{E}_{\mathbf{d}\in\mathcal{D}_{\mathbf{c}^t}}\left[\Delta H(\mathbf{l}|\mathbf{d},b^{t-1},\mathbf{w};\mathbf{c}\right]\eta\right)}{\sum_{\mathbf{c}'\in\mathcal{C}}\exp\left(\mathbb{E}_{\mathbf{d}\in\mathcal{D}_{\mathbf{c}}}\left[\Delta H(\mathbf{l}|\mathbf{d},b^{t-1},\mathbf{w};\mathbf{c}'\right]\eta\right)} \tag{5.15}$$

Positive values of $\rho \in [-\infty,\infty]$ encode a preference for focusing on areas where the learner should be most uncertain, $\rho = 0$ encodes random selection of local focus, and negative ρ encodes a preference for focusing on areas where the learner should be most certain.

For comparison, *Baseline* is a parameter-free model that assumed each intervention was a random draw from all possible interventions

$$P(\mathbf{c}^t) = \frac{1}{|\mathcal{C}|} \tag{5.16}$$

Global is a variant of the globally efficient intervention selection (Section 2) that attempted to select the globally most informative greedy test $\arg\max_{\mathbf{c}\in\mathcal{C}}\mathbb{E}_{\mathbf{d}\in\mathcal{D}_{\mathbf{c}}}\left[\Delta H(M|\mathbf{d},D^{t-1},\mathbf{w};C^{t-1},\mathbf{c})\right]$. It has one inverse temperature parameter $\theta \in [0,\infty]$ governing soft maximisation (Luce, 1959) over the global expected information gains. For this, we considered

$$P(\mathbf{c}^t|D^{t-1},\mathbf{w};C^{t-1}) = \frac{\exp(\mathbb{E}_{\mathbf{d}\in\mathcal{D}_{\mathbf{c}^t}}\left[\Delta H(M|\mathbf{d},D^{t-1},\mathbf{w};\mathbf{c}^t)\right]\theta)}{\sum_{\mathbf{c}\in\mathcal{C}}\exp(\mathbb{E}_{\mathbf{d}\in\mathcal{D}_{\mathbf{c}}}\left[\Delta H(M|\mathbf{d},D^{t-1},\mathbf{w};\mathbf{c})\right]\theta)} \tag{5.17}$$

As with the belief modelling, for Experiment 5 we marginalised over the unknown values of $\mathbf{w}$ rather than conditioning as in Experiments 1,2 and 4 as detailed in Appendix B.

Evaluation

All six models were fit to the data from all three experiments in the same way as the belief models. The results are detailed in Table 5.4.

Additionally, to compare model predictions of local focus choice $\mathbf{l}^t$ to participants' self reports in problem 7 in Experiment 5, we computed the likelihood of each local focus prediction on each test. This was done by calculating $P(\mathbf{c}|\mathbf{l},\eta,b^{t-1},\mathbf{w})$ for each of the local foci we considered, using a fixed common $\eta = 20$ to capture strong but non-deterministic preference for the most useful intervention(s). For each data point c^t, we then calculated which l^t assigned the most probability to c^t the intervention actually chosen by the participant. Figure 5.13 plots the most likely focus of participants' intervention choices in the final problem against the code assigned to their free responses.

Table 5.4: Intervention Model Fits

	Exps 1–2 (Bramley, Lagnado, & Speekenbrink, 2015)										
	Model	η M	η SD	ρ M	ρ SD	θ M	θ SD	N fit	M acc	logL	R^2	BIC
1	Baseline							3	0.38	-18262		36524
3	Edges	10.4	26	0.9	80			9	0.68	-14222	0.23	29338
2	**Effects**	**7.7**	**16**	**0.5**	**5**			**82**	**0.71**	**-10701**	**0.41**	**22296**
4	Confirmatory	3.9	74					9	0.43	-15368	0.14	31182
5	Mixed	15.1	152	0.7	33			32	0.66	-11145	0.39	23185
6	Global					6.1	4.1	4	0.85	-15619	0.14	31686
	Exp 4: Learning larger models											
	Model	η M	η SD	ρ M	ρ SD	θ M	θ SD	N fit	M acc	logL	R^2	BIC
1	Baseline							18	0.35	-32958		65917
2	Edges	9.3	138	25.9	1012			2	0.37	-28588	0.13	58196
3	Effects	6.2	39	1.5	6			31	0.59	-24213	0.27	49445
4	Confirmatory	3.8	8					24	0.49	-28721	0.13	57951
5	**Mixed**	**8.8**	**139**	**17.6**	**414**			**27**	**0.61**	**-23944**	**0.27**	**48907**
6	Global					4.9	4	18	0.66	-26652	0.19	53813
	Exp 5: Unknown strengths											
	Model	η M	η SD	ρ M	ρ SD	θ M	θ SD	N fit	M acc	logL	R^2	BIC
1	Baseline							14	0.35	-15365		30730
2	Edges	4.0	8	2.7	77			24	0.74	-13010	0.18	26850
3	Effects	3.2	9	2.9	218			7	0.52	-12992	0.14	26815
4	Confirmatory	2.5	14					12	0.46	-14180	0.04	28776
5	**Mixed**	**3.9**	**9**	**5.4**	**285**			**15**	**0.63**	**-12550**	**0.17**	**25931**
6	Global					3.0	8	39	0.70	-12850	0.14	26114

Note: Columns as in belief model (Table 5.3)

Results and discussion

The *mixed* local focus model was the best fitting model over the three experiments with the lowest total BIC, followed by *effects* then by the *global* focused model, then by *edges* and finally by *confirmation* and then baseline. However, there was a great deal of individual variation, suggesting that a single model does not capture the population well. More participants were best described by an *effects* focus (121) than a *mixed* focus (77), but each model received some support, with 58, 43, 36 and 35 individuals best fit by *global*, *confirmation* and *edge* focused and *baseline* models respectively. Additionally, the *effect* focus was the best fitting model overall in Bramley, Lagnado, and Speekenbrink (2015) where there was a strong tendency for participants to fix a single variable on at a time.

As Table 5.4 shows, *mixed* was the best overall fitting model for Experiment 4 and 5, and the majority of participants (277/370) were fit by one of the local uncertainty driven models. Furthermore, Figure 5.13 shows that for effect and edge queries, there was a strong correspondence between the most likely choice of focus l on Experiment 5 problem 7 and the coded explanation of that intervention's goal. This was not the

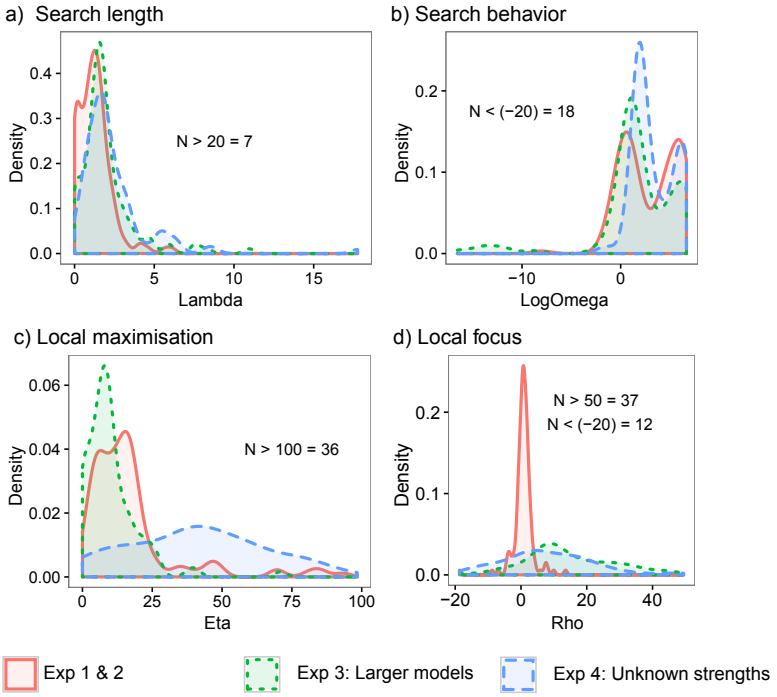


FIGURE 5.12: Gaussian kernel densities over fitted model parameters for all participants. a) Search length λ and b) log search behaviour $\log(\omega)$ according to *Neurath's ship* belief update model. c) Local maximisation parameter η and d) local focus choice parameter ρ under *mixed* local uncertainty based model. Since we report all participants' fits, there are some extreme values — poorly described by either model — that are not plotted. Annotations give the number of parameters above and below the range plotted.

case for tests where explanations were categorised as *confirmatory*. These were most frequently best described as effect focused tests of the root variable of the true model (labelled "X_1" in the plots).

As with the case of judgments, a moderate number of chance-level performing participants (35/370) were best described by the *Baseline* model. However, 58 participants across the three experiments were better described by the *Global*ly efficient testing model than any local testing models. However, these were not the highest performing participants in Experiment 5, with lower average scores than those described by the *edge* focused model. This suggests that we do not yet have a good model of these participants' choices.

5.9 General Discussion

Actively learning causal models is key to higher-level cognition and yet is radically intractable. We explored how people manage to identify causal models despite their limited

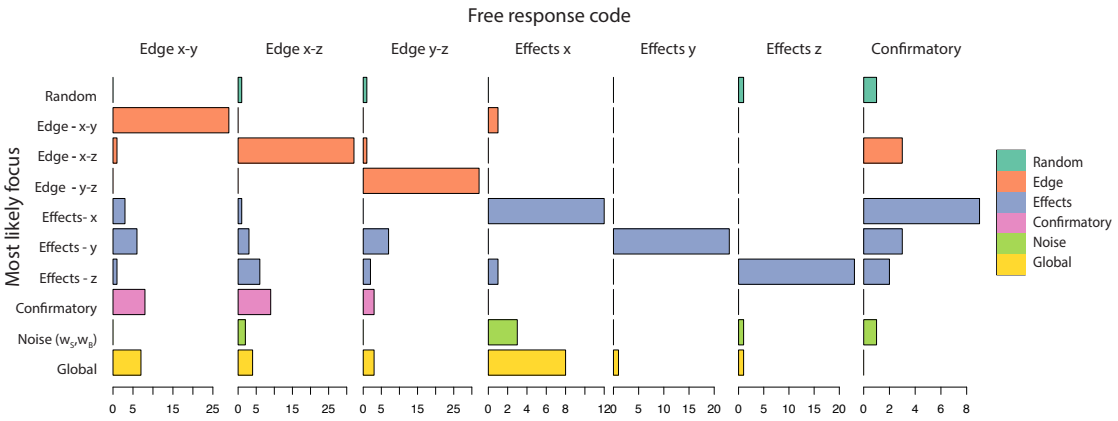


FIGURE 5.13: Model and free response correspondence. Each plot is for trials assigned a particular free response code, each bar is for the number of trials for which that local focus was most likely given the intervention choice. *Effect* and *edge* coded queries were also diagnosed as such by the model fitting while *confirmatory* coded queries were most likely to be diagnosed as querying the effects of the root node(s) in the true model which always was (or included) X_1.

computational resources. In three experiments, we found that participants' judgments somewhat reflected the true posterior, while exhibiting sequential dependencies. Further, participants' choices of interventions reflected average expected information, but were insufficiently reactive to the evidence that had already been observed and were consistent with being locally focused.

We could capture participants' judgment patterns by assuming that they maintained a single causal model rather than a full distribution. We proposed that participants considered local changes to improve the ability of their single model to explain the latest data and compared this account to two other proposals, one based on the idea that participants occasionally resample from the full posterior, and the other, a heuristic based on ignoring the possibility of indirect effects. While our *Neurath's ship* proposal fit best overall, all three proposals had merit, with *simple endorsement* winning out in Experiment 5 and more individuals better fit by *win-stay lose-sample*.

We captured participants' interventions by assuming they focused stochastically on different local aspects of the overall uncertainty and tried to resolve these, leading to behaviour that was comparatively invariant to the prior. Our modelling suggested a broad spread of local focuses both between and within participants.

By casting our modelling in the language of machine learning, we were able to make strong connections between our *Neurath's ship* model and established techniques for approximating distributions — sequential Monte-Carlo particle filtering and MCMC (specifically Gibbs) sampling. Likewise, we were able to explicate intervention selections

using the language of expected uncertainty reduction but relaxing the assumption that the goal was the reduction in global uncertainty in the full distribution. The combination of a single hypothesis (particle) and a Gibbs-esque search, nicely reflects the *Neurath's ship* intuition that theory change is necessarily piecemeal and that changes are evaluated against the backdrop of the rest of the existing theory.

5.9.1 Limitations of Neurath's ship

Like any theory, Neurath's ship was evaluated against a backdrop of a number of assumptions. We discuss some of these here.

Measurement effects

In order to explore incremental belief change it was necessary to elicit multiple judgments and to make two strong assumptions: (1) that these judgments reflected participants' true and latest beliefs; and (2) that the repeated elicitations did not fundamentally alter learning processes. To mitigate problems of these, we both incentivised participants to draw their best and latest guess at every time point during the tasks, and examined different reporting conditions to explore the influence of the elicitations on the learning process.

In Experiments 1 and 2 and the *remain* condition in Experiment 5, participants could leave parts of their hypothesis untouched if they did not want to change them. This had the strength of being minimally invasive; it did not push the learner to reconsider an edge that they would otherwise not have done merely because they have been asked about it again. However this came at the cost of conflating genuine incremental change in the learner's psychological representation with response laziness. To assuage this concern, in Experiment 4 and Experiment 5 *disappear*, we removed the participants' previous judgment after they had seen the outcome of the subsequent intervention, meaning that they would have to remember and re-report any edges they had previously judged (and not yet reconsidered). The slight reduction of dependence between *remain* and *disappear* conditions in Experiment 5, is consistent with the idea that being forced to re-report edges made it more likely that they would be reconsidered and potentially changed.

The *Neurath's ship* approach is related to anchor-and-adjust models (Einhorn & Hogarth, 1986; Petrov & Anderson, 2005) of sequential magnitude estimation. Hogarth and Einhorn found that, when mean estimates are repeatedly elicited from participants as

they see a sequence of numbers, the sequence of responses can be captured by a process whereby one stores a single value and adjusts it a portion of the way toward each new observed value. When judgments were elicited at the end of the sequence, participants behaved more like they had stored a subset of the values and averaged them at the end. In the same way, we can think of *Neurath's ship* as a process in which the current model acts as an anchor, and adjustments are made toward new data as it is observed. However, the higher complexity of causal inference, and the greater storage requirements for the individual episodes will presumably lead to greater pressure to use a sequential strategy rather than store. Arguably, step-by-step elicitation is a closer analogue to real-world causal inference than end-of-sequence because causal beliefs are presumably in frequent use while learning instances may be spread out, with no clear start or end.

Acyclicity

We adopted the directed acyclic graph as our model of causal representation here because it is a standard approach in the literature and is mathematically convenient. Furthermore, cyclic graphs were quite rare choices in Experiments 1 and 2 (where participants were permitted to draw them). Thus, we simply opted to rule them out in the instructions in later experiments.

However, in tasks where people draw causal models of real-world phenomena, they often draw cyclic or reciprocal relationships (Kim & Ahn, 2002; Nikolic & Lagnado, 2015), and many real world processes are characterised by bidirectional causality, such as supply and demand in economics or homoeostasis in biological systems. There are various ways to represent dynamic systems. One proposal is the dynamic Bayesian network (Dean & Kanazawa, 1989), which can be "unfolded" to form a regular acyclic network with causal influences passing forward through time. Another is the chain graph (Lauritzen & Richardson, 2002), in which undirected edges are mixed with directed edges and used to model the equilibria of the cyclic parts of the system.

Exploring these structures would require a change in the semantics of the experiment so that people could understand what they were reporting in the presence of dynamical interactions. However, given this, NS would offer a way of performing sequential, on-line, inference for such structures, using standard likelihood calculations for dynamic Bayes nets and chain graphs.

Evaluation of evidence

Another pragmatic limitation of the current modelling was the assumption of the noisy-OR functional form for the true underlying causal models. While we did take care to train participants on the sources of noise in all these Experiments, our own past work suggests that people may have simpler ways of evaluating how likely models would be to produce different patterns — for example, in Bramley, Dayan, and Lagnado (2015), we found participants' judgments could be captured by assuming they lumped sources of noise together and just counted the number of surprising outcomes under each model.

One possibility is that people actually formed likelihood estimates through simulation with an internal causal model. For instance, one might perform a mental intervention, activating a component of one's own internal causal model and keeping track of where the activation propagates. By simulating multiple times, a learner could estimate the likelihood of different outcomes under their current model (Hamrick et al., 2015), and by simulating under variations of the model, the learner could compare likelihoods generated on the fly. This simulation-based view provides a possible explanation for why participants more readily accommodated internal noise (e.g. $w_S \ll 1$) than background noise (e.g. $w_B \gg 0$). The former can be "built in" to the inferred connections in their model and reveal itself in mental simulation, while w_B is more of a mathematical "catch all" for all possible influences coming from outside the variables under focus (see Fernbach, Darlow, & Sloman, 2010, 2011, for similar proposals). The Neurath's ship perspective suggests that people lean on their surrounding network of assumptions about surrounding causes, controlling for these if they get in the way of local inference. By being omnipresent and affecting all the variables equally w_B was not possible to accommodate in this way.

Future experiments and modelling might relax the assumption of noisy-OR likelihoods and allow induction of more diverse functional forms (e.g. Lucas & Griffiths, 2010), or focus on well known domains where priors can be measured before the task. Another approach might be to render the noisy-OR formalisation more transparent by visualising the sources of exogenous noise alongside the target variables, for instance displaying varying numbers of nuisance background variables on screen for different background noise conditions.

Antifoundationalism

The core of Neurath's ship is the strong assumption that people consider only a single global hypothesis and make local changes within this. This is the "antifoundationalism" captured by Duhem–Quine thesis — any local theoretical claim is necessarily supported by surrounding assumptions. However, this may be too strong for some of the easier problems we considered here where the worlds may have been small and constrained enough for some people to reason at the global level. For the three variable problems in particular, some participants may have been able to consider alternatives at the level of the whole model, and thus able to shift from Fork to Chain etc with a single step.

While participants' judgments showed high sequential dependence, they did occasionally change their model abruptly. The theory of unexpected uncertainty (Yu & Dayan, 2003), and substantial work on changepoint tasks (Speekenbrink & Shanks, 2010) are associated with the notion that people will sometimes "start over" if they are having consistently poor predictions from their existing model. This relates to the idea, in philosophy of science, of a "paradigm shift" (Kuhn, 1962). The current *Neurath's ship* models do not naturally capture this but accommodate occasional large jumps by assuming a variable search length (k), meaning the search will sometimes be long enough to allow the learner to move to a radically different model in a single update. However we might also extend the *Neurath's ship* framework to include a threshold on prediction accuracy below which a learner will start afresh, for example by randomly sampling a model, or sampling from a hitherto unexplored part of the space. At present this is captured by the ϵ probability of sampling a new b^t at random on a given trial (which ranged between a probability of .03 in Experiment 5 and .2 in Experiment 4).

Selective memory

We assumed that participants' judgment updates were based on the recent data $\mathcal{D}_r^t$, collected since the last time they changed their hypothesis. This is quite frugal in the current context, as the learner rarely has to store more than a few tests' worth of evidence. It also captures the idea of semanticisation — that as one gradually absorbs episodic evidence into one's hypothesis, it becomes safe to forget it.

However, the particular choice of $\mathcal{D}_r^t$ is certainly a simplification. People may frequently remember evidence from before their latest change, and fail to store recent evidence, especially once their beliefs become settled. They might also collect summary evidence

at the level of individual edges, counting how often pairs of components activate together for example, or remember evidence about some components but not others, or only store evidence when it is surprising under the current model. In order to fit the models it was necessary to make simplifying assumptions that captured some form of halfway house between remembering everything and relying entirely on your hypothesis. Future studies might probe exactly what learners can remember during and after learning to get a finer-grained understanding of the trade-off between remembering evidence and absorbing it into beliefs.

Related to this, we fit a static search behaviour parameter to participants, finding evidence of moderate hill climbing. However, a more realistic depiction might be something more akin to simulated annealing (Hwang, 1988). Learners might begin searching with more exploratory moves $\omega \approx 1$ so as to explore the space broadly, and transition toward hill climbing $\omega = \infty$ as they start to choose what judgment to report. Alternatively, they might gradually reduce their search length k as pressure to settle on a model increases.

5.9.2 Alternative approximations and representations

The choice of Gibbs sampling, together with a single particle approximation, is just one of numerous possible models of structure inference. For example we found (data not shown) fairly good fits by replacing Gibbs sampling with a form of Metropolis-Hastings MCMC sampling — using an MC^3 proposal and acceptance distribution (Madigan & Raftery, 1994; Madigan, York, & Allard, 1995). The two approaches make similar behavioural predictions but differ somewhat in their internal architecture — a Metropolis-Hastings sampler would first generate a wholesale alternative to the current belief, then make an accept-reject decision about whether to accept this alternative, while the Gibbs sampler focuses on one subpart at a time and updates this conditional on the rest. Ultimately, the Gibbs sampler did a better job, helping justify the broader ideas of locality of inference implicit in the *Neurath's ship* proposal.

An interesting alternative approach to complex model induction via local computations (Fernbach & Sloman, 2009; Waldmann et al., 2008), comes from variational Bayes (Bishop, 2006; Weierstrass, 1902). The idea behind this is that one can simplify inference by replacing an intractable distribution, here the distribution over all possible models, with a simpler one which has degrees of freedom that can be used to allow it to fit as best as possible. A common choice of simpler distribution involves factorisation, with a multiplicative combination of a set of simpler parametrised distributions.

Thus, for causal inference one might make a mean-field approximation (Georges, Kotliar, Krauth, & Rozenberg, 1996) and suppose the true distribution over models factorises into independent distributions for each causal connection. Divergence between this approximation and the full model can then be minimised mathematically by updating each of the local distributions in turn (Jaakkola, 2001). This provides a different perspective on global inference based on local updates. Rather than a process of local search where only a single model is represented at any time, variational Bayes suggests people maintain many local distributions and try to minimise the inconsistencies between them. The biases induced by this process make the two approaches distinguishable in principle (Sanborn, 2015), meaning that an interesting avenue for future work may be to design experiments that distinguish between the two approaches to approximation in cognition. The truth in our case may be somewhere in between. For instance, in the current work, we assumed people were able to use recent evidence to estimate their local uncertainty conditional on the rest of the structure, and thus choose where to focus interventions. To the extent that learners really represent their beliefs with lots of local uncertainties, their representation becomes increasingly variational.

5.9.3 Choosing interventions aboard Neurath's ship

The largest difference in intervention choices between experiments was that in Experiment 5 constrained interventions (e.g. $\text{Do}[X_1 = 1, X_2 = 0]$) were chosen much more frequently. One explanation for this is that participants might have been forced to focus their attention more narrowly in Experiment 5, to compensate for their additional uncertainty about the noise by using more focused testing. Another possibility is that the different subject pools drove this difference. It is possible that mTurk's older and educationally diverse participants (Experiments 1 and 2) gathered evidence differently from the young scientifically trained UCL undergraduates (Experiment 5). This might have driven the tendency toward more tightly constrained tests in Experiment 5.

The idea that people relied on asking a mixture of different types of locally focused question, was borne out by our analysis of the coding of participants' free explanations. Explanations almost always focused on one specific aspect of the problem, most frequently on a particular causal connection, or what a particular component can affect, but also sometimes on parameter uncertainty or, on later tests, confirming their current hypothesis. Furthermore, participants almost always referred to a mix of different

local query types over the course of their six tests. The apparent shift toward confirmatory testing on the last trial is sensible, since participants knew they would not have more tests to follow up anything new they might discover. Indeed, this shift would be normative in various settings.

Subjective explanations are notoriously problematic (Ericsson & Simon, 1980, 1993; Russo, Johnson, & Stephens, 1989). Therefore, we must be careful in interpreting these results. One common issue is that eliciting responses concurrently with performing a task can change behaviour, invalidating conclusions about the original behaviour. We minimised this issue by eliciting explanations just after each intervention was chosen, before its outcome was revealed. Additionally, we did not find any difference in the distribution of interventions on the free response trials and those chosen the first time participants identified the Chain structure.

A second issue is that there are limits on the kinds of processes people can describe effectively in natural language, with rule based explanations being typically easier to express than those involving more complex statistical weighting and averaging. That is, even if someone weighed several factors in coming to a decision, they might explain this by mentioning only the most significant, or recently considered of these factors, falsely appearing to have relied on a one-reason decision strategy. There is an active debate about this, including suggestions that people's explanations for their choices are, in general, post-hoc rationalisations rather than genuine descriptions of process (Dennett, 1991; Johansson, Hall, Sikström, & Olsson, 2005), but also refutations of this interpretation (Newell & Shanks, 2014).

In sum, taken with appropriate caution, we suggest that this analysis does provide a valuable window on participants' subjective sense of their active testing, with their relatively specific focus on one aspect of the uncertainty at a time consistent with the idea that they rely on a mixture of heuristic questions.

The models pinned down interventions less tightly than beliefs in the sense that there was a great deal of spread in the individuals best fit across the models, and the proportional reductions in BIC were smaller. There are various possible reasons for this. Firstly, the models of belief change generally predicted one or few likely models, whereas there are typically many interventions of roughly equal informativeness to an ideal learner (see Figure 5.2), which could be performed in many different orders. This sets the bar for predictability for interventions much lower than for the causal judgments.

Secondly, to the extent that learners chose interventions based on a reduced encoding of the hypothesis space, we are also forced to average over our additional uncertainty about exactly which hypotheses or alternatives they were considering at the moment of choice (Markant & Gureckis, 2010).

A third issue is that of whether and how learners represented current uncertainty, and recruited this in choosing what to focus on. In the current work we assumed that learners were somewhat able to track the current local uncertainties and use these to choose what to target next. The modelling revealed that, relative to the local intervention schema, the majority of participants did tend to focus on the areas of high current uncertainty (shown by the predominantly positive ρ in Figure 5.12 d) but we do not yet have a model for how they did this. It is plausible that learners used a heuristic to estimate their local confidence. For example, a simple option would be to accrue confidence in an edge, (or analogously in the descendants of a variable or in the current hypothesis) for every search step for which it is considered and remains unchanged, reducing confidence every time it changes. In this way confidence in locales that survive more data and search become stronger, approximately mimicking reduction in local uncertainty.

We considered just three of a multitude of possible choices of local focus. These encompass most extant proposals for human search heuristics, encapsulating modular (Markant et al., 2015) constraint seeking (Ruggeri & Lombrozo, 2014) and confirmatory (Klayman & Ha, 1989) testing, placing all three within a unified schema and also showing that many learners dynamically switch between them.

Participants' free responses provided a complementary perspective, suggesting that even initial tests were generated as solutions to uncertainty about some specific subpart of the overall uncertainty space — often the descendants of some particular variable or the presence of some particular connection. This suggests that the most important step in an intervention selection may not be the final choice of action but the prior choice of what to focus on next. This is captured in our model, under which the values of different interventions for a chosen focus do not depend on D^{t-1}. This means learners need not do extensive prospective calculation on every test but can learn gradually, for instance through experience and preplay (Pfeiffer & Foster, 2013), which interventions are likely to be informative relative to generic types of local focus. This knowledge could then be transferred to subsequent tests, and translated to tests with different targets — e.g. if $Do[X_1 = 1]$ is effective for identifying the effects of X_1 then $Do[X_2 = 1]$ will be effective for identifying the effects of X_2.

It is worth noting from these data that even when participants' interventions were relatively uninformative from the perspective of ideal or even our heuristic learners, their explanations would generally reveal that they *were* informative with respect to some other question or source of uncertainty. For example, participants' tests that were uninformative with respect to identifying structure were often revealed, through our free response coding, to have been motivated by a desire to reduce uncertainty about internal w_S or background w_B noise.[14] From this perspective we might think of even the completely uninformative intervention choices — e.g. fixing all the variables — as legitimate tests of illegitimate hypotheses — e.g. hypotheses that were outside of the space of possibilities we intended participants to consider — such as whether fixed variables actually always took the states they were fixed to. More research is needed to explicate these internal steps leading up to an active learning action, but the implication based on the current research is that the solution will not require that the learner evaluate all possible outcomes of all possible actions under all possible models, but rather reflect a mixture of heuristics that can guide the gradual improvement of the learner's current theory.

5.9.4 The navy of one

At the start we argued that our Neurath's ship model could be seen as a single particle combined with an MCMC search. As such, we are claiming *Neurath's ship* as a form of boundedly rational approximate Bayesian inference. However, it is important to consider the point at which an approximation becomes so degenerate that it is merely a complicated way to describe a simple heuristic. Many would argue that this line is crossed long before reaching particle filters containing a single particle, or Markov chains lasting only 1 or 2 steps. It is certainly a leap to claim that such a process is calculating a proper posterior.

One alternative to starting from a normative computational level account and accepting a distant algorithmic approximation, is to start from the algorithm, i.e., the simple rules, and consider a computational account such as satisficing (Simon, 1982) that provides adequate license. Our account shares two important problems with this, but avoids two others.

[14] We might have extended the computational model of Bayesian inference to incorporate joint inference over models and parameters which would have incorporated this aspect of testing. However, this would have complicated analyses since participants were ultimately only incentivised to identify the right connections.

One shared problem is the provenance of the rules - i.e., the situation-specific heuristics. We saw this in the manifold choice of local foci for the choice of intervention — we do not have an account of whence these hail. This is a common problem in the context of the adaptive toolbox (Gigerenzer, 2001) — it is hard to have a theory of the collection of tools.

A second shared problem follows on from this - namely how to choose which rule to apply under which circumstance. In our case, this is evident again in the mixtures of local focus rules — we were not able to provide a satisfying account of how participants make their selection of focus on a particular trial. The metaproblem of choosing the correct heuristic is again a common issue for satisficing approaches.

By contrast with a toolbox approach, though, our account smoothly captures varying degrees of sophistication between individuals. For instance, with the *Take the best* heuristic, Gigerenzer et al. (1999) give an attractive description of one-reason decision making that often outperforms regression in describing people's decisions from multiple cues. However, subsequent analyses have revealed that participants behave somewhere between the two (Newell & Shanks, 2003; Parpart et al., in revision) often using more than one cue, but certainly less than all the information available. Thus, to understand their processing we must be able to express the halfway houses between ideal and overly simplistic processing (Lieder & Griffiths, 2015). In the same way, the approximate Bayesian perspective allows us to express different levels of approximation lying between fully probabilistic and fully heuristic processing, with the simplest form of Neurath's ship lying at the heuristic end of this road.

A further benefit of our account is the ease of generalisation between tasks. Heuristic models are typically designed for, and are competent at, specific paradigms. Since they lack a more formal relationship with approximate rationality, they are hard to combine or often to apply in different or broader circumstances.

Here, we assumed that learners made updates at the level of individual directed edges. Again this is just one illustrative choice, but our model is consistent with the idea that the learners altered beliefs by making changes local to arbitrary sub-spaces of an unmanageable learning problem. We showed that so long as the learner's updates are conditioned on the rest of their model, and are appropriately balanced, the connection to approximate Bayesian inference can be maintained through the ideas of MCMC sampling and a single-particle particle filter. A sophisticated learner might be able to update

several edges of their causal model at a single time, with a more complex proposal distribution. However, on a larger scale this is still likely to be a small subset of all potential relata that learner has encountered, meaning even the most sophisticated learner must lean on their broader beliefs for support.

In lower level cognition, inference takes place over simple quantities like magnitudes and is certainly probabilistic in the sense that humans can achieve near optimal integration of noisy signals in a variety of tasks including estimation (Miyazaki, Nozaki, & Nakajima, 2005) and motor control (e.g. Körding & Wolpert, 2004). At the top end of higher level cognition we have a global world-view, and explicit reasoning characterised by its single track nature. Rather than claiming these are completely different processes (Evans, 2003), the approximate probabilistic inference perspective can accommodate the whole continuum. At the lower level the brain can average over many values, as in particle filtering (Abbott & Griffiths, 2011), with a whole fleet of Neurath's ships, or via lots of long chains (Gershman et al., 2012; Lieder et al., 2012). In higher level cognition, however, the hypothesis space becomes increasingly unwieldy, and inference becomes increasingly approximate as it must rely on smaller fleets, i.e., fewer hypotheses, and more local alterations in the face of evidence. At the very top we have a navy of one, grappling with a single global model that can only be updated incrementally. It is worth noting that individuals can then play the role of particles again in group behaviour (Courville & Daw, 2007), giving us approximate inference all the way up.

In sum, retaining the Bayesian machinery is valuable even as it becomes degenerate, because it allows us to express heuristic behaviour without resorting to separate process models or abandoning close connections to an appropriate computational level understanding.

5.9.5 Scope of the theory

We modelled causal belief change as a process of gradually updating a single representation through local, conditional edits. While we chose to focus on causal structure inference within the causal Bayes net framework here, there is no reason why this approach should be limited to this domain. By taking the *Neurath's ship* metaphor to reveal an intuitive answer as to how people sidestep the intractability of rational theory formation (van Rooij et al., 2014), we can start to build more realistic models of how people generate the theories that they do and how and why they get stuck. We might explain the induction and adaptation of many of the rich representations utilised in cognition

by analogous processes. Future work could explore the piecemeal induction of models involving multinomial, continuous (Nodelman, Shelton, & Koller, 2002; Pacer & Griffiths, 2011) or latent variables (Lucas, Holstein, & Kemp, 2014); unrestricted functional forms (Griffiths, Lucas, Williams, & Kalish, 2009); hierarchical organisation (Griffiths & Tenenbaum, 2009; Williamson & Gabbay, 2005); and temporal (Pacer & Griffiths, 2012) and spatial (Battaglia et al., 2013; Ullman et al., 2012; Ullman, Stuhlmüller, Goodman, & Tenenbaum, 2014) semantics. One possibility is the combination of production rules (Goodman, Tenenbaum, Feldman, & Griffiths, 2008) and local search to model discovery of new hypotheses in situations where the space of possibilities is theoretically infinite. The sequential conditional re-evaluation process illustrated by our *Neurath's ship* model shows how this radical antifoundationalism need not be fatal for theory building in general.

5.10 Conclusions

This chapter proposed a new model of causal theory change, based on an old idea from philosophy of science — that learners cannot maintain a distribution over all possible beliefs, and so must rely on sequential local changes to a single representation when updating beliefs to incorporate new evidence. It showed that we can provide a good account of participants' sequences of judgments in three experiments and argued that our model offers a flexible candidate for explaining how complex representations can be formed in cognition. We also analysed participants' information-gathering behaviour, finding it consistent with the thesis that learners focus on resolving manageable areas of local uncertainty rather than global uncertainty, showing cognisance of their learning limitations. Together these accounts show how people manage to construct rich, causally-structured representations through their interactions with a complex noisy world.

Chapter 6

The role of time in causal learning

> "*For all the points of the compass, there's only one direction and time is its only measure.*"
>
> — TOM STOPPARD

Research on human causal learning has predominantly focused on learning from trial-by-trial covariation between variables based on observations of the system (Cheng, 1997; Deverett & Kemp, 2012; Gopnik et al., 2001; Perales & Shanks, 2007), and on active interventions on the system (Bramley, Lagnado, & Speekenbrink, 2015; Meder et al., 2014; Sloman & Lagnado, 2005; Steyvers et al., 2003). However, people utilise a range of sources of information in causal learning (Lagnado et al., 2007) and human causal knowledge goes beyond mere expectations about covariation (Gerstenberg et al., 2015; Sloman & Lagnado, 2015). To be able to predict and diagnose causality in real-world situations, human causal knowledge must often include beliefs about *how long* different relationships take to work (Griffiths & Tenenbaum, 2009; Lagnado & Sloman, 2006) — for example we know that turning on the heating activates the boiler almost instantly, but that the boiler will take a few minutes to make a radiator hot, and the radiator will take much longer to heat the room. Expectations about causal delays can in turn support structure inference (Buehner & McGregor, 2006; Griffiths & Tenenbaum, 2009; Hagmayer & Waldmann, 2002; Kemp, Goodman, & Tenenbaum, 2010) because the consistency between an observed event stream and the predictions of different causal models provides evidence about the underlying relationships. This temporal information is unavailable to a purely contingency-based learner. This Chapter focuses on the role of time in causal structure induction.

A strength of the CBN framework is that by defining a language for expressing possible models it allows causal learning to be framed as a Bayesian model induction problem, where the learner uses observed evidence to infer an underlying causal structure. From this perspective, people are generally found to be effective causal learners who make inferences that are broadly normative (e.g. Griffiths & Tenenbaum, 2005; Lagnado & Sloman, 2002, 2004, 2006; Steyvers et al., 2003) but also exhibit the signatures of various inductive biases and cognitive limitations (Bramley, Dayan, & Lagnado, 2015; Bramley, Lagnado, & Speekenbrink, 2015; Coenen et al., 2015; Mayrhofer & Waldmann, 2016; Rottman & Hastie, 2014, 2016).

Strengths aside, a shortcoming of the CBN framework is that Bayesian networks do not naturally encode the temporal or spatial dimensions of causal beliefs, and so say nothing about their role in causal inference (cf. Gerstenberg et al., 2015; Gerstenberg & Tenenbaum, to appear; Goodman, Tenenbaum, & Gerstenberg, 2015; Wolff & Shepard, 2013). Consequentially, many studies have focused on situations where information about time and space is non-diagnostic or abstracted away. When temporal cues have been pitted against statistical cues experimentally, judgments have tended to be dominated by temporal information (Burns & McCormack, 2009; Frosch et al., 2012; Lagnado & Sloman, 2004, 2006; Schlottmann, 1999). Furthermore, when researchers have tried to instruct participants to ignore event timing, participants still treated the observed timings of events to be diagnostic (McCormack et al., 2016; White, 2006b). These results suggest that people have strong assumptions about time's role in causality (see also Bechlivanidis & Lagnado, 2013, showing that the influence runs both ways). In the current chapter we take a novel approach: we eliminate statistical contingency information, and focus exclusively on what participants can learn about causal structure from temporal information alone.

Structure of the chapter

The structure of this chapter is as follows. First, we review the literature on causal learning and time. After describing the learning problem we focus on, we then outline the Bayesian framework and methodology used to explore human causal learning in time. Experiment 6 explores one-shot causal structure judgments, based on a single observation of a simple device operating through time. In Experiments 7 and 8, we look at how people integrate evidence from multiple clips of the same device. In Experiment 8, we focus on temporal delays in a situation where there is no temporal order cue: three

events always occur in the same order, but the variability and correlation between the timings of the events is either more consistent with a chain or a fork structure. Finally, we discuss the scope of our findings and propose future research.

6.1 Existing research

Temporal information is relevant for causal judgments in at least two ways. The *temporal order* of events is important since causes cannot precede their effects. The *timing* of events provide additional information, since they are diagnostic about the true underlying causal structure when multiple possible structures are consistent with the observed temporal order.

Temporal order

The assumption that causes precede effects is at the heart of our notion of causality (Hume, 1740) meaning that, *prima facie*, the order in which events occur is a highly important cue to causality. Inference from perceived order appears to be natural, almost automatic. For example, Wolff and Shepard (2013) cite multiple reports, following a 1997 power blackout in New York, of people having the sensation that an action they had taken just before the blackout (touching a doorknob, plugging in an appliance, jamming a ceiling fan) was its cause. Magicians use this to trick their audiences into believing they can affect objects at a distance, snapping their fingers just before revealing their masterstroke (Kuhn, Caffaratti, Teszka, & Rensink, 2014).

Precedence also forms the basis for many legal judgments, with establishment of the order of the events in a case often playing a large role in attribution of responsibility for a crime (Lagnado, 2011; Lagnado & Gerstenberg, in press). Additionally, an important concept in economics, *Granger causality* (Granger, 1969), uses the extent to which past values of one variable can be used to predict current variation in another as a marker for causation.

Rottman and Keil (2012) explored causal induction in situations where variables were measured at discrete intervals. For example, one might measure barometric pressure and precipitation on successive days. Finding that barometric pressure was high on Monday and Friday and it rained on Tuesday and Saturday invites the inference that high pressure causes rain. In seven experiments, the authors find that people readily

attribute causal relationships from variables that changed state at time $t-1$ to those that changed state at time t, and do so even when a cover story suggests there should be sequential independence. They argue that people's default representation of causality is as a qualitatively ordered sequence of changes, and suggest that estimating statistical dependence across multiple independent instances, as in contingency-driven structure learning, is a more difficult, less natural mode of causal reasoning.

Experienced event order also affects people's causal judgments when events take place in continuous rather than discretised time. Lagnado and Sloman (2006) explored a situation that contrasted trial-by-trial covariation with temporal order cues. In their experiment, a virus propagates through a computer network causing computers to be infected with different temporal delays. Participants' task was to infer the structure of these computer networks based on having observed the virus spreading through the network multiple times. Participants preferred causal models that matched the experienced order in which computers got infected, even when trial-by-trial covariation cues went against temporal order cues.

Several studies further suggest that people are reluctant to endorse causal connections between events which appear to occur at the same time. Burns and McCormack (2009) found that by age 6 to 7, children strongly favour a $B \leftarrow A \rightarrow C$ common cause over a $A \rightarrow B \rightarrow C$ chain when they observe B and C happening simultaneously and after A. Even when the causal mechanism is plausibly instantaneous (as in Frosch et al., 2012; Lagnado & Sloman, 2006; McCormack et al., 2016), people tend to attribute simultaneous activations of components to a common cause. However, previous work has not looked at what people infer from situations in which observed simultaneity cannot be attributed to a common cause, but must either be instantaneous or coincidental. Additionally, little work has looked specifically at how people integrate evidence of events occurring in different orders over multiple trials (although see Lagnado & Sloman, 2006).[1]

Event timings

Going beyond temporal order, we can also consider the exact timing of events as another source of information about causal relationships (Hagmayer & Waldmann, 2002). Using *only* temporal precedence to guide judgments would put everything that ever happened

[1] We use the $\rightarrow$ operator to denote a causal relationship between events (e.g. $A \rightarrow B$ means A caused B), and $\succ$ operator to denote event *order* (e.g. $A \succ B$ means A preceded B).

on equal footing as a candidate cause; a switch you switched a year ago would be just as likely a cause of a light turning on, as one you just switched.[2] In this section we discuss what role event timing plays in guiding causal judgments.

In the associative tradition, causal relationships are treated as another form of learned association, where the constant conjunction, and temporal as well as spatial contiguity between two variables naturally leads to their being associated by the cognitive system (Hume, 1748). Since increased intervals rapidly reduce the rate of associative learning (Grice, 1948; Shanks & Dickinson, 1987; Wolfe, 1921), associative theories generally predict that judgments of causality will show this same pattern. Making similar predictions, early cognitive theories (Ahn et al., 1995; Einhorn & Hogarth, 1986) suggested that the more distant two events are in time, the more costly it will be to sustain the first event in working memory long enough to relate it to the second event, leading to monotonic reduction in causal judgments. Lagnado and Speekenbrink (2010) identify an additional normative reason for why delays will often lead to reduced judgments of causality. All things being equal, the longer the gap between putative cause and effect, the more likely it is that other events may have occurred in the meantime that could have also caused the effect.

Humans are able to make causal inferences that are sensitive to expectations about event timing. When participants are given information about causal mechanisms that imply different delays, their resultant causal judgments are strongly influenced by expectations about average delay length and variability of the mechanisms. Seeing shorter-than- as well as longer-than-expected intervals leads to reduced judgments of causality (Buehner, Cheng, & Clifford, 2003; Buehner & May, 2002, 2004; Greville & Buehner, 2010; Hagmayer & Waldmann, 2002; Schlottmann, 1999). For example, seeing a regular light bulb come on several seconds after switching a switch is rated as less causal. However, the case is different if you learn that it is an energy saving bulb which takes time to warm up.

As well as unexpected time intervals, variability in intervals across trials has been shown to reduce judgments (Greville & Buehner, 2010; Greville, Cassar, Johansen, & Buehner, 2013; Lagnado & Speekenbrink, 2010). However, these studies have focused on situations in which there is a single candidate cause–effect pair. In this chapter, we explore the more general problem of inferring the causal structure of multiple variables based on observations of events in time.

[2]Of course, there are more intervening events in the former case, providing another possible avenue capturing why the latter is a better candidate (e.g. Lagnado & Speekenbrink, 2010).

Neuroimaging data also support the idea that timing expectations play a role in causal learning (Jocham et al., 2016). In two behavioural experiments, participants' task was to identify occasional rewarding events in event streams. The results showed that both associative and "contingent" — or theory-dependent — learning take place simultaneously and in separable brain circuits — the former predominantly in the amygdala, and the latter in the orbitofrontal cortex. Amygdala learning was associative in the sense that it learned relations between rewards and preceding events irrespective of task instructions. Orbitofrontal learning was contingent in the sense that it depended dynamically on instructions about what delays to expect for genuine stimuli and reward relationships, only attributing rewards to appropriately timed stimuli. A central goal in causal learning research is to understand where these theory-dependent judgments come from. How are people and animals often able to make strong and sensible "one-shot" inferences about causal structure, without explicit instruction, in situations where naive statistical learning algorithms would require much more data?

Several researchers have suggested that causal theories might underpin such "one-shot" inferences (Goodman et al., 2011; Griffiths, 2005; Griffiths & Tenenbaum, 2009; Kemp, Goodman, & Tenenbaum, 2010). The idea is broadly that, over the course of development, people organise their causal general knowledge hierarchically, with the core abstract features of causation at the top and increasingly domain- and context-specific features below. Each level of the theory generates a probability distribution on variables at the level below, and the more specific the subdomain, the greater the constraints on the space of possible hypotheses. As a learner's world knowledge gets richer, their causal judgments can rely more strongly on identification of the right domain and application of domain-specific knowledge and constraints, resulting in apparent "one-shot" inferences (see also Lake et al., 2015).

The theory-based causal inference framework provides an explanation for the role of temporal expectations in causal induction. By learning the typical cause–effect delays in a particular domain, a learner can use this knowledge to rapidly identify new connections when candidate events are appropriately spaced in time. Griffiths (2005) showed how different expectations about delay distributions allow for strong one-shot inferences about a causal process. In his experiments, participants made causal judgments about "nitroX" barrels that were causally connected and exploded in different sequences. Because different causal models imply very different event timings, the Bayesian model was able to rapidly infer the causal structure from an observed sequence of exploding barrels. Building on this work, Pacer and Griffiths (2012) model causal influence in situations

where a discrete event affects the rate of occurrence of another variable in continuous time. In particular, they capture people's judgments about the causal strength of variables that affect the rate of bacteria death in a population over a number of days (cf. Greville & Buehner, 2007). Extending this approach, Pacer and Griffiths (2015) also capture inferences about relationships for which the influence of the cause on the effect is expected to last for some time before it gradually dissipates. Using this model, Pacer and Griffiths explain participants' inferences about which of three occasionally occurring seismic waves affected the rate of occurrence of earthquakes (see Lagnado & Speekenbrink, 2010, Experiment 7). As predicted by the model, participants' judgments were affected by uncertainty about the number of intervening events rather than the absolute intervals between putative causes and effects.

Pacer and Griffiths' approach is well-suited to capturing situations where events alter the *rate* of occurrence of other events. However, it does not readily apply to situations in which causes bring about their effects exactly once. For example, an event at A in their model might increase the *number of activations* of B that you expect to see over the next 5 seconds from 0.1 to 1.1. However in their representation the number of events that occur in total is inherently stochastic. This means that the occurrence of the cause might sometimes result in no activation of B, or in several activations of B. In this chapter, we are interested in situations in which the causal relation between two events is singular — that is, the cause affects the effect exactly once.

In summary, research has established that temporal information plays an important role in how people make causal judgments (Lagnado & Sloman, 2006; Rottman & Keil, 2012; Sloman, 2005). Causal inference seems to be driven by temporal information partly via automatic (Michotte, 1946/1963) and developmentally basic (Burns & McCormack, 2009) mechanisms, but also through more complex theory-contingent modes of thinking (Griffiths & Tenenbaum, 2005). While previous work has explored representation of causality in time (Griffiths, 2005; Pacer & Griffiths, 2012), no research to date has proposed a model that is sensitive to temporal order and incorporates expectations about intervals between particular events.

6.2 Modelling causal induction from temporal information

Despite the wealth of research on time and causal learning, temporal information has not been subsumed into a unifying framework for understanding how causal beliefs are

formed to the same extent as contingency information has. In this section, we lay out a learning problem that isolates the role of temporal information. We then present our Bayesian approach to modelling learning in this situation, distinguishing learning based on information about *temporal order* alone from learning based on forming parametric expectations about *temporal intervals* between causes and effects.

6.2.1 The learning problem

To isolate temporal information, we focus on situations where the learner must identify the causal structure of a system made up of a number of components that are causally related but in which the causal links take time to propagate. We assume that the causal relationships are known to be generative and sufficient in the sense that the activation of a cause component will invariably lead to the activation of its effect component(s), but where the delays between activation of the cause and the effect are variable across instances. In Experiments 6 and 7 we focus on judgments about the causal structure of a simple system with two causal components A and B and an effect component E that form a hypothesis space of seven possibilities (Figure 6.1). In Experiment 8, we will focus on a more restricted space with a single cause component S and two components A and B that are either its direct or indirect effects. Evidence in all the experiments consists of clips that show how the different components of a causal device activate over time. In Experiments 1 to 2 participants are told that the parentless components in the device activate due to background causes while in Experiments 3 the learner activates the system themselves.

The different causal connections of a causal structure might exhibit different delays — for example, in the A-fork (see Figure 6.1) it might take longer on average for A to cause B than for A to cause E. Furthermore, the same connection might also exhibit variability in delays across trials — for example, A's causing E might be subject to longer or shorter delays on different occasions. As a consequence of this variability, many causal structures can generate several qualitatively different *orders* of activation.

6.2.2 Bayesian models of learning

From a Bayesian perspective, learning is the process of updating a probability distribution over the true state of the world, where the ground truth is treated as a random

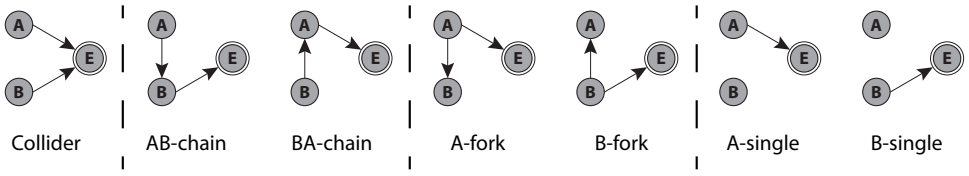


FIGURE 6.1: Possible causal structures in Experiments 6 to 8. The arrows indicate the direction of the causal relationship. Dotted lines indicate different types of structure. *Note*: the Collider is conjunctive — both A and B must occur for E to occur.

variable and its possible values make up the hypothesis space. A Bayesian learner updates their prior probability distribution into a posterior distribution as evidence is observed. The posterior from one learning instance becomes the prior for the next, and this process continues as evidence is received. With sufficient evidence, the learner's subjective beliefs eventually approximate the ground truth provided that the hypotheses are distinguishable and the hypothesis space contains the ground truth.

Exact Bayesian inference is intractable for most realistically complex problems. However, for a suitably constrained problem space like the one explored here, Bayesian inference provides a powerful framework for understanding human learning. We can look at how people update their beliefs as evidence is presented, and learn about the prior assumptions they bring to the task.

In the current context, the random variable we are interested in is the true underlying causal structure $m \in \mathcal{M}$ in the set of possible structure hypotheses $\mathcal{M}$, and data will take the form of n observed patterns of component activations over time $\mathbf{d} = (d_1, d_2, \ldots, d_n)$. We update a prior belief about the possible underlying structures $P(M)$ to a posterior belief over the structures given the data $P(M|\mathbf{d})$ using Bayes theorem

$$P(M|\mathbf{d}) \propto p(\mathbf{d}|M) \cdot P(M), \tag{6.1}$$

where $p(\mathbf{d}|M)$ is the likelihood function over structures $\mathcal{M}$.

For inference to proceed, the learner needs a likelihood function determining how likely each structure would be to exhibit the set of experienced temporal patterns $\mathbf{d}$. We first propose a class of models based on simple likelihood functions that ignore the exact timing of events but assign likelihoods simply based on their temporal ordering. We consider two models that differ in whether they allow for instantaneous causation, that is, causes and effects happening at the same time. We then consider a richer framework that incorporates expectations about causal delays. We show how, based on the principles of Bayesian Ockham's razor (MacKay, 2003), both approaches form

preferences for different causal structures requiring neither contingency information nor pre-existing expectations about the duration or variability of the delays.

6.2.3 Only order matters

Likelihood functions

The order of events constrains what structures are capable of having produced the observed evidence. We capture the information contained in the temporal order of events in a simple model that divides its likelihood evenly across all order-consistent patterns. Hence, any particular sequence of component activations has likelihood $1/N$, where N is the number of distinct temporal orderings consistent with that structure (Figure 6.2b and c, columns). In the following, we use the $\succ$ operator to denote event order. For example, $A \succ B \succ E$ means that A preceded B which preceded E. $AB \succ E$ means that A and B happened simultaneously before they were succeeded by E.

In the A-fork, A is the cause of both B and E, therefore this structure is consistent with patterns in which A preceded both B and E ($A \succ B \succ E$, $A \succ E \succ B$ and $A \succ BE$, see Figure 6.2a) but inconsistent with any pattern where either B or E precede A. Whether $AB \succ E$ or $AE \succ B$ are consistent with the A-fork depends on whether one assumes causes and effects can occur simultaneously. In order to test whether people make this assumption, we will compare two variants of our model to participants' judgments. *Order non-simultaneous* (Order_N) makes the non-simultaneity assumption meaning only events that strictly precede other events are candidate causes. *Order simultaneous* (Order_S) relaxes this assumption, such that an event can be the cause of another event even if they occur at the same time. For Order_N, the AB-chain is only consistent with $A \succ B \succ E$ (Figure 6.2b, second column). For Order_S, the AB-chain is also consistent with $AB \succ E$ and $A \succ BE$, and thus this model variant spreads its likelihood more widely.[3]

Because some structures are compatible with fewer kinds of evidence patterns than others, the order models will tend to favour them over a more flexible structure that can

[3]Note that several additional possible patterns are not pictured: $AE \succ B$, $BE \succ A$, and ABE. We do not use these in our experiments because, by being inconsistent with all structures under the non-simultaneity assumption their appearance would force a simultaneity assumption on participants. However, if people actually make the simultaneity assumption, $AE \succ B$ and $BE \succ A$ are each consistent with one fork and one single, and ABE is consistent with all seven structures. Thus we divide the Order_S likelihoods across these additional structures, yielding columnwise likelihoods of $[\frac{1}{6}, \frac{1}{4}, \frac{1}{4}, \frac{1}{6}, \frac{1}{6}, \frac{1}{8}, \frac{1}{8}]$. We excluded any patterns in which E alone occurred first, since we instructed participants that E is always caused by either A, B, or both.

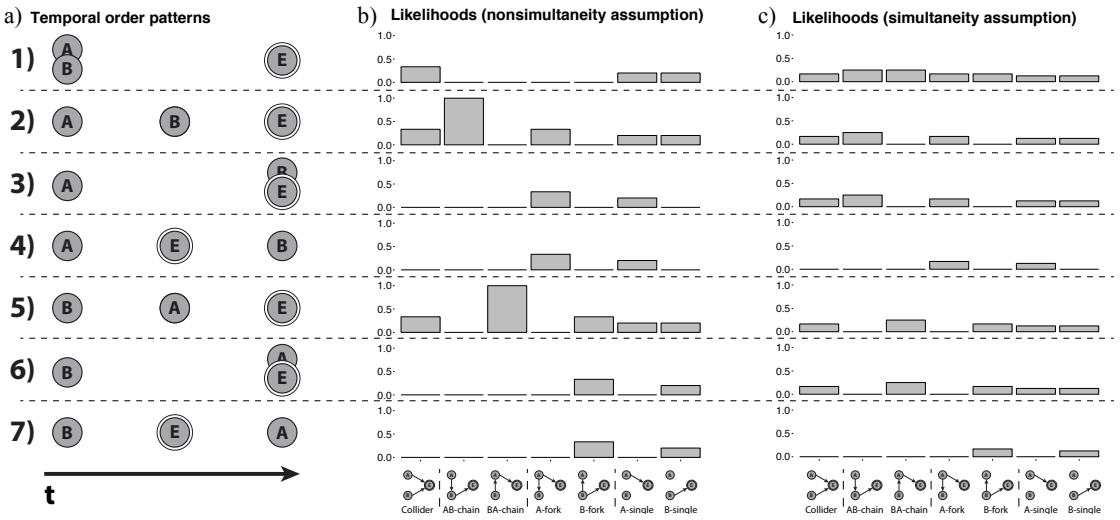


FIGURE 6.2: a) Seven possible qualitative temporal patterns of three events A, B, and E. Likelihood functions for the pattern types given the seven different causal structures with non-simultaneity assumption b) or simultaneity assumption c).

also produce the evidence seen. For example, under the non-simultaneity assumption, pattern $A \succ B \succ E$ in row 2) is the only pattern consistent with the AB-chain, while A-single is consistent with all but two types of patterns and thus spreads its likelihood much more widely. Switching focus from Figure 6.2's columns to its rows gives a perspective on the models' posterior predictions. For instance, upon observing a device that activates in the $A \succ B \succ E$ order, Order_N will favour the AB-chain, even though it has not ruled out the Collider, the A-fork, or either of the two single-link structures A-single and B-single.

As another example, after observing pattern 1) $AB \succ E$, the Order_N model will rule out all structures except for the Collider, A-single, and B-single. Between these remaining structures, it prefers the Collider since it is consistent with fewer types of pattern. In contrast, the Order_S model cannot rule out any structure based on this evidence. It has a slight preference for the AB-chain and the BA-chain since these two structures are compatible with the fewest number of different temporal order patterns.

Inference

After seeing data $\mathbf{d}$ in the form of one or several temporal order patterns, inference proceeds by updating a prior over causal structures $\mathcal{M}$ to incorporate these data.

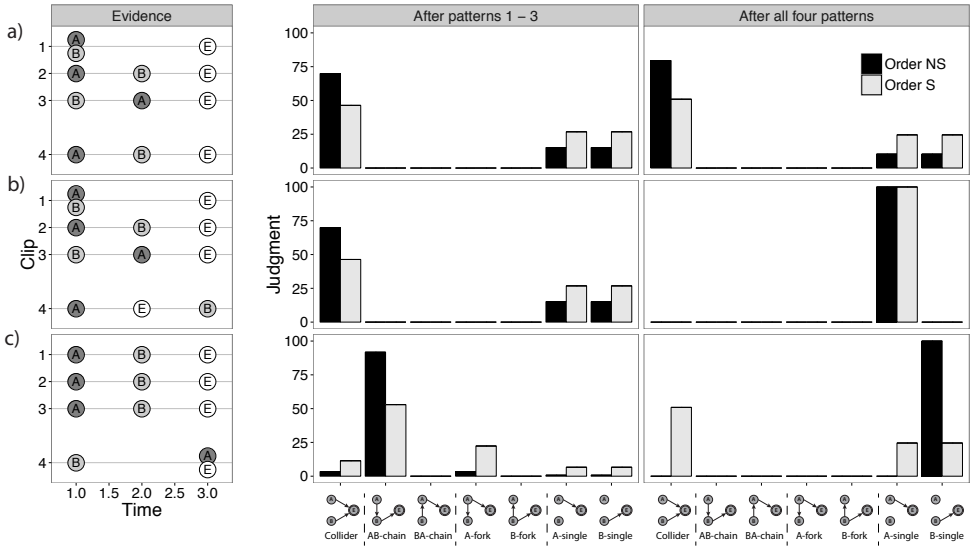


FIGURE 6.3: Three examples of order model predictions. Left hand side: Sets of 4 time series showing staggered activation of components A, B and E. Right hand side, model posteriors after seeing clips 1–3 (left column), and after having seen all four patterns (right column).

The order models only consider the qualitative ordering of the component activations, for example $\mathbf{d} = (d_1 = \{A \succ B \succ E\}, d_2 = \{AB \succ E\}, \ldots)$, where d_i indexes independent observations of the device. The models yield various posterior beliefs based on different sequences of temporal activation $\mathbf{d}$. For example, starting from a uniform prior over the seven structures, Figure 6.3 shows posteriors under the simultaneous and non-simultaneous assumptions based on having observed three patterns of activation d_1, d_2, and d_3, and then again after having observed a fourth pattern d_4. In the first example (top row), both non-simultaneous and simultaneous models favour the Collider after d_1, d_2, and d_3 and their preference increases with d_4. In the second example (middle row), both models prefer the Collider after d_1, d_2, and d_3 but switch upon d_4 which rules out out all the structures except the A-single. In the third example (bottom row), the two assumptions lead to quite different predictions, with the non-simultaneous model preferring the Collider and the simultaneous model preferring the B-single after seeing all the clips.

6.2.4 Timing matters

Generative model

The order models make the strong assumption that any activation pattern whose temporal order is consistent with the device is equally likely. While simple to work with,

the assumption is inconsistent with more specific beliefs about delays between causes and effects. For example, people may believe that causes take a certain amount of time to bring about their effects and that these delays will be similar across instances. To capture these intuitions, we need richer models than Order_N and Order_S – models that incorporate assumptions about how the events and their timings are being brought about.

In our task, an observed temporal pattern d_i consists of the activation times t_X of the three components A, B, and E; thus $d_i = \{t^i_A, t^i_B, t^i_E\}$. We will use t_{XY} to refer to the temporal interval between the activations of X and Y (i.e. $t_{XY} = t_Y - t_X$). Additionally, we will use $t_{X\to Y}$ to distinguish causal delays from temporal intervals t_{XY} which are not necessarily causal.

Independent causes We start with formalising the timing of independent causes which do not have any parents in the causal structure (such as variables A and B in the Collider, A-single, and B-single). Analogously to CBNs, in which independent causes are assumed to be statistically independent of each other (i.e. uncorrelated), we define independent causes to be temporally independent of each other as well as independent from the (artificially determined) beginning of the clip. The natural candidate for modelling such events is the exponential distribution, which is "memory-less". This property means that how long you expect to wait for an event is independent of how long you have already been waiting for it. Thus, the information that another (independent) event has happened does not alter your expectation about the time of the next event. If X is an independent (i.e. parentless) cause then the timing of X is determined by

$$p(t_X|\lambda) = \lambda e^{-\lambda t_X} \tag{6.2}$$

with $p(t_X|\lambda) = 0$ for activation times smaller than 0 and expectation $\frac{1}{\lambda}$.

Causal links The generalisation of the exponential distribution is the gamma distribution. It introduces time dependence, and it is therefore the natural candidate to model the relative timing of causally related events. Gamma distributions can be defined by a shape parameter α and an expectation μ. Under the assumption that X causes Y, the timing of Y depends upon the timing of X such that $t_Y = t_X + t_{X\to Y}$ with $t_{X\to Y}$ being gamma distributed:

$$p(t_{X\to Y}|\alpha,\mu) = \frac{(\frac{\alpha}{\mu})^{\alpha}}{\Gamma(\alpha)}(t_{X\to Y})^{\alpha-1}e^{-\frac{\alpha}{\mu}t_{X\to Y}} \tag{6.3}$$

with $p(t_{X\to Y}|\alpha,\mu) = 0$ for temporal delays smaller than 0 (i.e. no backwards causation).

Figure 6.4 shows examples of gamma distributions for different parameter values. The gamma distribution is flexible and allows to represent a continua of short (small μ) to long (large μ) and variable (low α) to reliable (high α) delays.

As $\alpha \to \infty$, the gamma distribution becomes increasingly centred around its expected value, capturing what we will call "positive" time dependence (e.g. Figure 6.4, solid and dashed lines). One's expectation about the time of an effect increases following the observation of its cause, peaking around its mean and then dropping away again. For $\alpha = 1$ the gamma distribution is an exponential distribution. Values of $\alpha < 1$ capture "negative" dependence whereby upon observing a cause one expects to see its effect either right away or in a very long time (e.g. Figure 6.4, dot-dashed line).

Colliders/Common-effect structures Within this framework, the Collider (i.e. common-effect structure) presents a special modelling challenge since it involves a joint influence of two distinct causes. There are various plausible combination functions for capturing this kind of joint influence. We explicitly stipulate in all experiments that the Collider structure is conjunctive, meaning that the activation of E occurs only after the activations of both A and B and, by implication, the arrival of both of their causal influences at E. To model this, we consider the t_E in a Collider structure to be the maximum of the two unknown causal delays for $t_{A\to E}$ and $t_{B\to E}$ offset by their activation time

$$t_E = \max[t_A + t_{A\to E}, t_B + t_{B\to E}] \tag{6.4}$$

with $t_{A\to E}$ and $t_{B\to E}$ being gamma distributed (see Equation 6.3) and t_A and t_B being exponentially distributed events (see Equation 6.2).[4] Note that a *disjunctive* Collider is modelled by simply using the minimum instead of the maximum in Equation 6.4 (see Equations B.4 and B.6 in Appendix B).

[4]We derive the full equations for the Collider likelihood assuming shared parameters for the input connections (as in Delay_P) and separate parameters (as in Delay_I), in Appendix B.

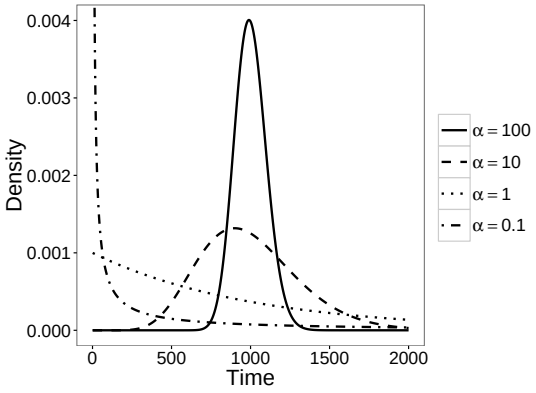


FIGURE 6.4: Four example gamma distributions. All have a mean of $\mu = 1000$ ms but differ in their shape α. The exponential distribution is the case where $\alpha = 1$.

Likelihood functions

The generative model laid out above provides the formal tools we need to determine the likelihood of any observed temporal pattern given a structure hypothesis. To distinguish different causal structures, we translate the absolute timings of a set of events into specific cause–effect pairings, depending on the parents $\text{pa}(X)$ of each variable under the structure at hand. For instance, absolute timings $\{t_A, t_B, t_E\}$ will be translated into $\{t_A, t_{AB}, t_{BE}\}$ with $t_{AB} = t_B - t_A$ and $t_{BE} = t_E - t_B$ under the AB-chain hypothesis. Dependent on different beliefs about the underlying causal structure and delay distributions, the same set of observed activation times will be more or less likely as we will illustrate below.

Sometimes it may be reasonable to assume that the different connections in a causal system have the same underlying delay distribution (e.g. they might all be components of the same type). In other situations, we might expect completely different delays for different parts of a process (for example it might take millions of years for the wind to wear through a rock face but only seconds for the freed rock to fall and cause a landslide). We can embody these different assumptions with different model variants. The *pooled model* (Delay_P, Figure 6.5a) has a single α and μ parameter for all the delays within a single structure $m \in \mathcal{M}$. In contrast, the *independent model* (Delay_I, Figure 6.5c) has separate parameters α_e and μ_e for each causal connection $e \in E_m$ where E_m is the list of all edges in structure m. To capture weaker assumptions (e.g. that the delay distributions for relationships within a device are related but not identical), one could extend this with a *hierarchical model* (Delay_H, Figure 6.5b) that combines expectations about the variability of the different distributions within a device via hyperparameters that define distributions for α and μ, although we do not do this here.

We start by describing the likelihood function of the *pooled* Delay$_P$ variant of the model. The likelihood of a temporal pattern d_i given a causal structure $m \in \mathcal{M}$ with timings governed by parameters λ, α and μ, is the product of the likelihoods of the relative delays between causes and effects that result from mapping the absolute event timings $t_X \in d_i$ onto the structure of the model

$$p(\mathbf{d}|\lambda, \alpha, \mu; m) = \prod_{i \in 1:n} p(d_i|\lambda, \alpha, \mu; m) = \prod_{i \in 1:n} \prod_{t_X^i \in d_i} p(t_X^i - t_{\mathrm{pa}(X)}^i|\lambda, \alpha, \mu; m) \tag{6.5}$$

with $p(t_X^i - t_{\mathrm{pa}(X)}^i|\lambda, \alpha, \mu; m)$ being either gamma or exponentially distributed (see Equation 6.3 and Equation 6.2, respectively) depending on whether X has a parent or not (and assuming that the structure is not a Collider).

For the Collider, we have to determine the joint likelihood of t_{AE} and t_{BE}. Note that we use "$\rightarrow$"s to distinguish the unknown true delays from the observed inter-event intervals since either t_{AE} or t_{BE} may include some time spent "waiting" for the other causal influence to arrive. Observed event timings depict one of two mutually exclusive cases: either the causal influence of A arrived later (i.e. $t_{A\rightarrow E} = t_{AE}$) overshadowing the timing of B's causal influence (i.e. $t_{B\rightarrow E} \leq t_{BE}$) or the causal influence of B arrived later (i.e. $t_{B\rightarrow E} = t_{BE}$) overshadowing the influence of A (i.e. $t_{A\rightarrow E} \leq t_{AE}$). The joint likelihood of the observed intervals are then given by the sum of their individual likelihoods

$$p(t_{AE}, t_{BE}|\alpha, \mu) = p(t_{AE}|\alpha, \mu) \cdot p(t_{B\rightarrow E} \leq t_{BE}|\alpha, \mu) + p(t_{BE}|\alpha, \mu) \cdot p(t_{A\rightarrow E} \leq t_{AE}|\alpha, \mu) \tag{6.6}$$

with $p(t_{AE}|\alpha, \mu)$ and $p(t_{BE}|\alpha, \mu)$ being gamma distributed (see Equation 6.3), and $p(t_{A\rightarrow E} \leq t_{AE}|\alpha, \mu)$ and $p(t_{B\rightarrow E} \leq t_{BE}|\alpha, \mu)$ following the cumulative distribution function of the gamma distribution (i.e. the integral over Equation 6.3 with upper bound t_{AE} or t_{BE}, respectively; see Appendix for a more detailed derivation).

In the general case, λ, α, and μ are unknown. To get the (marginal) likelihood of the data given the structure, which is our target for Equation 6.1, we have to marginalise out the parameters by integration, assuming some prior distribution over λ, α, and μ[5]

[5]Concretely, we used an Exponential(0.1) prior for α, an Exponential(0.0001) prior on μ and an Exponential(10000) prior on λ, corresponding to a weak expectation for positive dependence, shorter delays and frequently occurring independent causes.

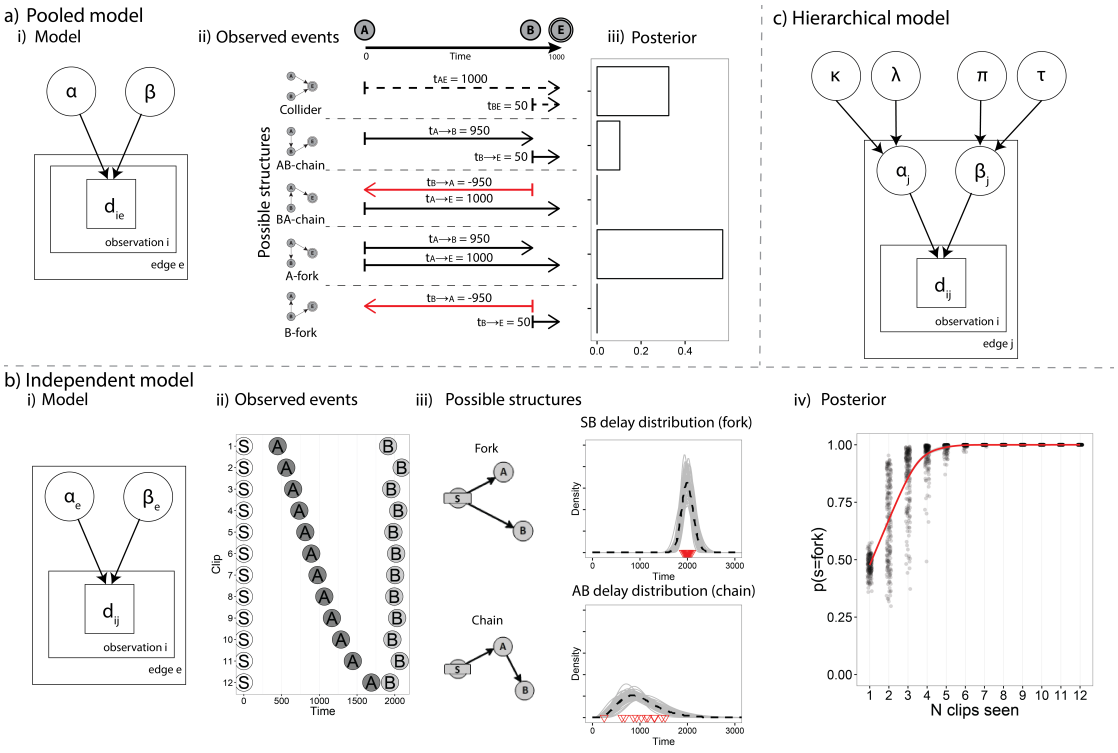


FIGURE 6.5: Delay sensitive models and predictions. a) i. *Pooled* Delay_P in plate notation. ii. Example of inference in the pooled model. Observed event timings are mapped onto causal delays under different models. Each row shows the causal delays assuming a different structure. For the Collider, dashed lines indicate that one or other causal delay may be shorter than the observed intervals. Red arrows indicate structures that can be ruled out based on order alone. iii. Posterior predictions of the delay model assuming priors of $M \sim \text{Unif}(\frac{1}{7})$, $\alpha \sim \text{Exp}(0.1)$, and $\mu \sim \text{Exp}(0.0001)$. b) i. *Independent* Delay_I model in plate notation. ii. 12 patterns of evidence. iii. Posterior marginal inference for two possible structures. The plots show posterior delay samples (grey lines) and their overall density (dotted black line). Both structures share the same $t_{S\to A}$ delays, but the high variance of t_{AB} relative to t_{SB} means these data were more likely produced by a fork as shown in iv., which plots the posterior probability of the fork structure averaged over subsets of the 12 clips (red line gives smoothed average, black dots give posteriors for samples). c) An example of a *hierarchical* Delay_H model in plate notation, where different components have different distributions but are related by hyperparameters.

$$p(\mathbf{d}|m) = \int p(\mathbf{d}, \lambda, \alpha, \mu|m)\ d\lambda\ d\alpha\ d\mu \tag{6.7}$$

$$= \int p(\mathbf{d}|\lambda, \alpha, \mu; m) \cdot p(\lambda, \alpha, \mu|m)\ d\lambda\ d\alpha\ d\mu \tag{6.8}$$

$$= \int p(\mathbf{d}|\lambda, \alpha, \mu; m) \cdot p(\lambda|s) \cdot p(\alpha|m) \cdot p(\mu|m)\ d\lambda\ d\alpha\ d\mu \tag{6.9}$$

We discuss how we approximated these integrals and sensitivity to priors in Appendix B.

To see how this timing sensitivity supports causal structure inferences, let us assume

that a learner observed the following order of activation: $A \succ B \succ E$. If they make the Delay_P assumption that cause–effect delays for the connections in this device come from the same distribution, we would expect their belief about whether the underlying causal structure was a Collider, an AB-chain, or an A-fork to shift depending on t_{AB} and t_{BE}. Intuitively, if t_{AB} and t_{BE} are similar, this seems most consistent with an AB-chain. However, if t_{BE} is very small this seems more consistent with the A-fork (in which $t_{A\to B}$ and $t_{A\to E}$ would be similar). If t_{AB} is very small then the device might be a Collider (where we would expect $t_{A\to E}$ and $t_{B\to E}$ to be similar). Delay_P makes these predictions via Bayesian Occam's razor. Essentially, it assumes all causal delays of the connections in a device follow the same underlying gamma distribution. Even if we have only a vague idea what specific form this distribution takes (as specified by α and μ), the model will still tend to favour whatever causal hypothesis renders these causal event timings the most similar on average. The more tightly clustered the inferred delays are, the more compact the generative causal delay distribution can be (here a high average α parameter), which leads to higher likelihoods assigned to the data points. See Figure 6.5a for an illustration of this point.

Inference in the *independent* Delay_I model (Figure 6.5c) proceeds in same way, but with separate parameters for the delay distributions of the different causal connections $c \in C$ [e.g. $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{|C|})$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_{|C|})$]. That is, it assumes there is no relationship between the delays of different parts of a causal device. The distribution of delays implied by mapping event timings onto different causal models can still be diagnostic, provided one interacts with the same device more than once. Figure 6.5c gives an illustration of this. Here, the temporal intervals t_{SB} are consistently around 2 s, while t_{SA} and t_{AB} are much more variable. We can explain these patterns of evidence more parsimoniously by assuming that the true structure is an S-fork with a regular $S \to B$ connection and an irregular $S \to A$ connection. It is not impossible that the true structure is a chain, but the chain structure cannot explain the additional systematicity in the data whereby the $t_{S\to A}$ and $t_{A\to B}$ intervals almost perfectly cancel out.[6]

6.2.5 Summary

In summary, the non-simultaneous and simultaneous order models (Order_N and Order_S) operationalise inference based purely on the qualitative ordering of observed activations.

[6] We note though that with additional assumptions about the functioning of the device the reverse inference might hold. For example, if the $A \to B$ connection was somehow designed to cancel out variation in $S \to A$ so as to lead to a reliable t_B.

These models show how certain structures can be ruled out, and some of the remaining structures preferred, based on the order of events alone. What structures are ruled out depends on whether simultaneous events are considered consistent with causation. While these order-based models are good at ruling out inconsistent causal structures, they are limited in their ability to distinguish between structures that are consistent with the observed order of events.

The delay-based models *pooled* Delay_P, *independent* Delay_I and *hierarchical* Delay_H make inferences within the space of hypotheses not-yet-ruled-out by Order_N, but distribute their likelihood very differently depending on the expected rate and variability of the various inter-event intervals. Assuming an uninformative prior on shape α and mean μ, the pooled delay model Delay_P favours whichever structures render the experienced inter-event intervals the most regular across all connections and all instances, while the independent delay model Delay_I favours whichever structures imply the most regular within-edge delays on average, even if these differ considerably for different connections. A hierarchical model would make predictions somewhere in between, allowing that different connections can have different delays but that they are still related.

6.3 Overview of Experiments

6.3.1 The task

We designed a task environment in which participants observed causal devices exhibiting one or several patterns of activation, and then made judgments about how they thought the components of that device were causally connected. Evidence was presented in the form of short movie clips. Each clip simply showed three components, (A, B, and E in Experiments 6 and 7, and S, A and B in Experiment 8), which were represented by circles and arranged in a triangle (see Figure 6.6 bottom left). During each clip, all three components activated by turning from white to grey (Experiments 6 and 7) or from white to yellow (Experiment 8). Activated components remained coloured until the end of the clip. To minimise people's context-specific expectations about what causal structures or delays were more likely a priori, we kept the task abstract. Participants were not told anything about what kinds of causal processes underlie the activation of the different components.

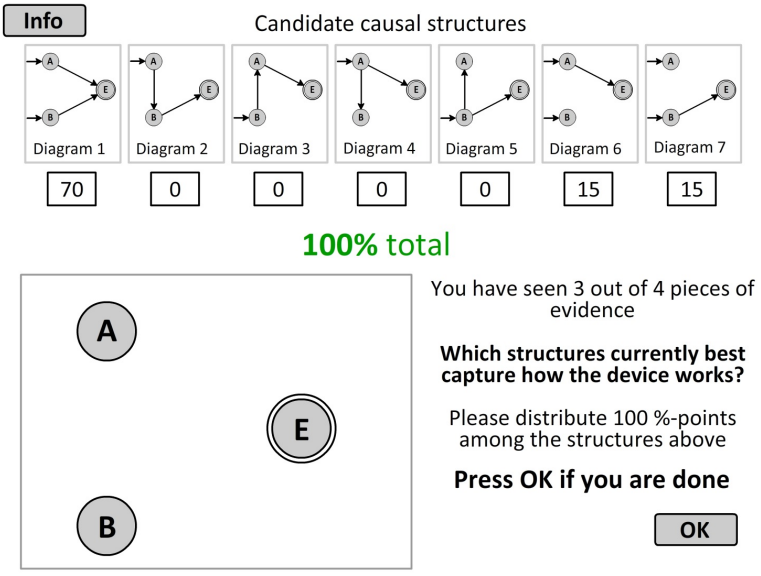


FIGURE 6.6: The experiment interface for Experiments 6–7. Clips are shown in the bottom left panel and judgments elicited at the top.

Possible causal structures

As discussed in the introduction, we restricted the space of possible causal structures to seven in Experiments 6 and 7 (see Figure 6.1) and two in Experiment 8. In Experiments 6 and 7, each structure featured two candidate causes A and B and one effect E. Participants were informed that the Collider structure is conjunctive, meaning that both A and B must activate in order for E to occur. In Experiment 8 there was a starting component S and two candidate effects A and B. The true structure was either a chain (e.g. $S \rightarrow A \rightarrow B$) or a fork (e.g. $A \leftarrow S \rightarrow B$).

Eliciting judgments

In order to have a fine-grained measure of participants' beliefs, we asked participants to distribute 100 percentage points over the set of possible candidate causal structures, such that each value indicated their belief that the given structure is the one that generated the observed evidence (see Figure 6.6 top). We can then directly compare participants' distributions over the structures with the predicted posterior distributions based on our different models.

6.4 Experiment 6: One-shot inferences

In Experiment 6 we explored one-shot inference. We asked participants to make judgments about causal devices after watching a single clip and replaying it several times. We varied the timing and order of the activation of the three components systematically across problems. Depending on whether or not participants rule out instantaneous causation, we expected the judgments to better match the predictions of the non-simultaneous or simultaneous order model, respectively. If participants' judgments were, in addition to temporal order, also sensitive to timings, we expected them to assign more points to structures that imply similar cause–effect delays (e.g. a fork if B occurs very early as in clip 2 shown in Figure 6.7a, and a Collider if B occurs very late as in clip 6).

6.4.1 Methods

Participants and materials

Thirty-one participants (18 female, $M_{age} = 36.8$, $SD_{age} = 11.9$), recruited from Amazon Mechanical Turk, took part in Experiment 6. The task took 15 minutes (SD = 8.7) on average and participants were paid at a rate of $6 an hour. The task interface was programmed in Adobe Flash 5.5.[7] Demos of all three experiments are available at `http://www.ucl.ac.uk/lagnado-lab/el/nbt`.

Stimuli and model predictions

Participants made judgments about nine devices in total. For each device they saw evidence in the form of a single, replay-able video clip. All clips began with a 500 ms interval after which the first component(s) activated. The clip then lasted another 1000 ms whereupon the final component(s) activated. We chose a range of clips in which A occurred at the start and E at the end, varying where B fell in between the two (see Figure 6.7a, clips 1–7), and then two clips in which E occurred earlier than B (clips 8 and 9). We obtained model predictions by computing the posterior for $P(M|\mathbf{d})$ for Order_N and Order_S, and Delay_P, assuming learners began each problem with a uniform prior

[7] Flash has been shown to be a reliable way of running time-sensitive experiments online (Reimers & Stewart, 2015). We checked the time-accuracy of our code during development finding it highly accurate.

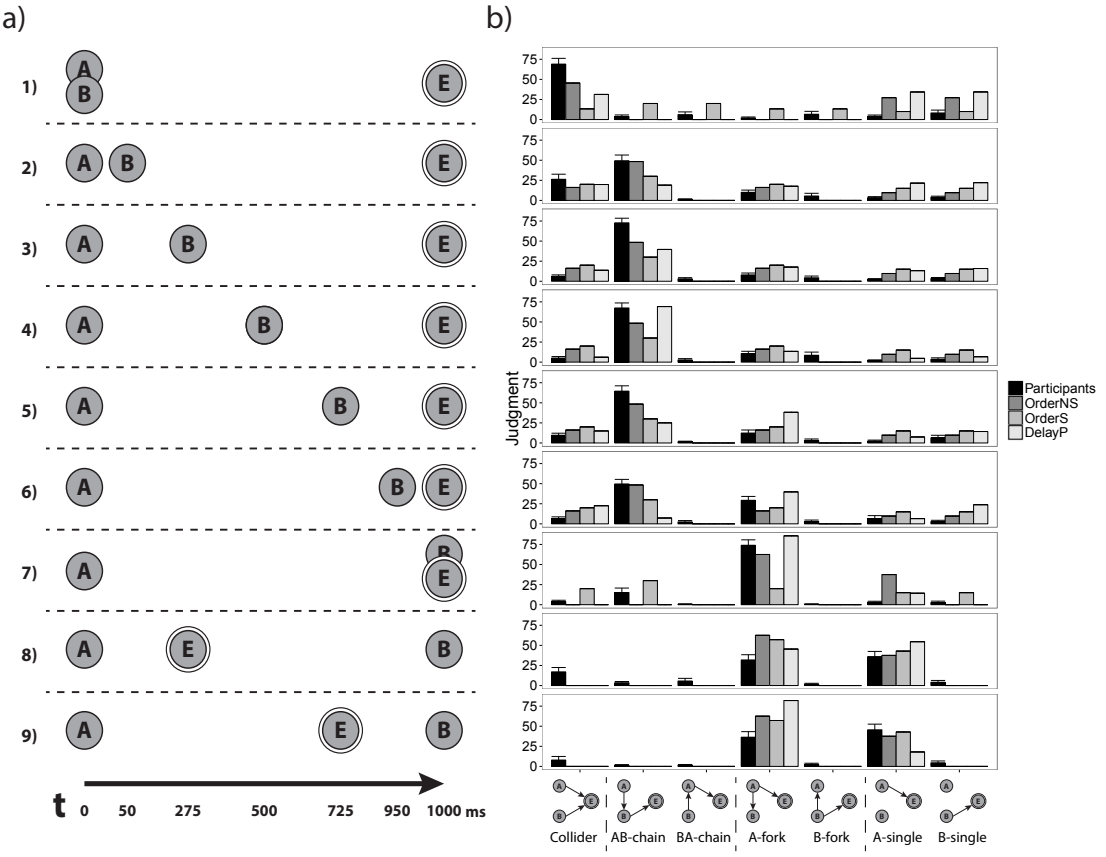


FIGURE 6.7: a) The timeline for each clip type in Experiment 6 b) Participants' averaged judgments after viewing each clip (black bars) and predictions by the different models (grey bars). Error bars show standard errors.

across structures and the base rate λ, and a diffuse prior over causal delay parameters α and μ (see Appendix).[8]

Order$_N$ and Order$_S$ model predictions do not vary across clips 2-6 where order was always $A \succ B \succ E$ with both models favouring the AB-chain but Order$_N$ having a stronger preference (see Figure 6.7b).

They also do not differentiate between clips 8 and 9 (both $A \succ E \succ B$) where both model variants slightly favour the A-fork. Order$_S$ predicts a broad spread across structures for clips 1 and 7, in both cases slightly favouring a chain structure while the Order$_N$ favours the Collider and A-fork, respectively. Sensitivity to timing leads to predictions that differ across clips 2 to 6. Delay$_P$ favours the Collider and A-single and B-single structures when B occurs relatively early, and prefers the chain when B occurs relatively late. Delay$_P$ is also sensitive to the difference in the timing of E between clips 8 and 9, preferring the

[8]Note that Delay$_I$ does not make predictions here since it requires repeated evidence to form preferences about the connections.

Table 6.1: Experiment 6: Order and Delay Models Compared to Participants' Judgments

Model	r	r_s	Mode match	RMSE	N con (N dis)
Baseline				20.1	0
Order_N	0.90	0.76	78%	11.1	12 (10)
Order_S	0.71	0.75	56%	16.4	1 (1)
Delay_P	0.80	0.64	44%	15.8	3 (4)

Note: Model fits assuming the Collider was conjunctive. r = average Pearson's r correlation between average assignments to structures within each device and model predictions. r_s = average Spearman's rank correlation within problems. Mode match = proportion of problems where participants' modal choice matched model's. RMSE = root mean squared error. N = Number of individuals best correlated by model (con= assuming conjunctive Collider, dis= assuming disjunctive Collider).

A-fork if E happens relatively late and the A-single if it occurs early. Finally, it puts more probability mass on the two single structures than the other models.

Procedure

In the instructions, participants were familiarised with the seven causal structure diagrams, and the response format. Participants then completed the 9 problems in random order. Components A and B were counterbalanced such that on approximately half of the problems faced by each participant their roles were reversed (e.g. B would occur at the start rather than A and their responses flipped for analysis). In each trial, participants observed a single clip of a device and then replayed that same clip. After the fourth replay, participants distributed 100 percentage points across the 7 possible devices displayed at the top of the screen. They were allowed to replay the clip a fifth and final time before finalising this judgment and moving on to the next device. Participants could only move on if their indicated answers summed to 100%. The causal devices were displayed at the top of the screen in the same order for all problems. For half of the participants, the order of the seven devices was as depicted in Figure 6.6 while for the other half it was reversed.

6.4.2 Results

There was no effect of counterbalancing on participants' judgments, with no interactions between the A-B counterbalance and participants' assignment of percentage points across the structures, nor with order in which the structures were presented on the screen.

As Figure 6.7b and Table 6.1 show, the Order_N model captures participants' judgments best overall here. Comparing participants' responses directly with model predictions,

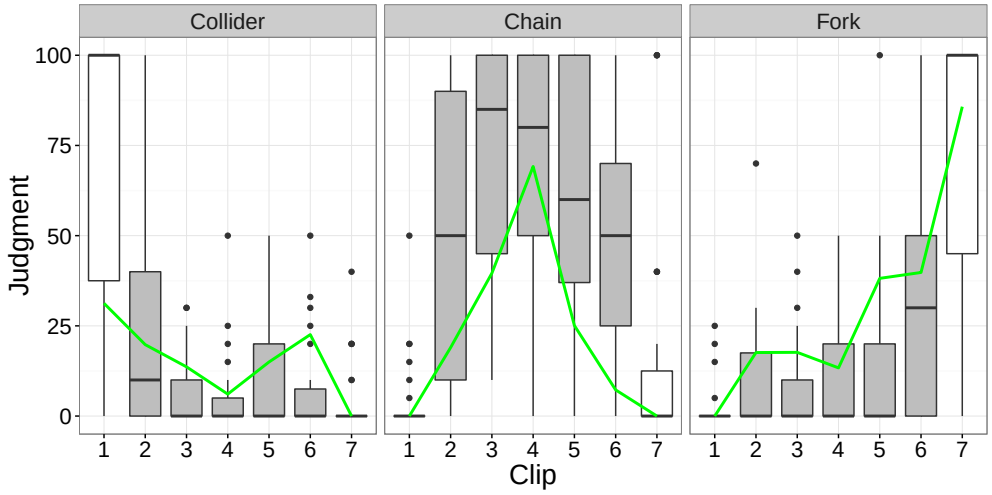


FIGURE 6.8: Comparison of probability assignments to Fork, Chain and Collider structures for clips 1–7 (cf. Figure 6.7), in which B appears at 0, 50, 275, 500, 725 950 and 1000 ms after A, with E always occurring at 1000 ms. Boxplots show participants' median and upper and lower quartiles, participants with judgments $\pm > 1.5$ interquartile range are plotted separately. Results in text are relative to the six middle bars (grey). Green lines denote Delay_P model predictions.

we see that, on average, judgments were well correlated with the Order_N, more so than for Order_S, and Delay_P. While Delay_P beats Order_S in terms of Pearson's correlation r, it is a little worse at getting participants' rank order right as shown by the lower Spearman correlation r_s.

As we see in Figure 6.7b participants assigned some mass to the Collider for clips 8 and 9, suggesting that some participants forgot or disregarded our instruction to think of the Collider as conjunctive (i.e. both causes were needed to generate the effect). To check this we also computed model predictions assuming a disjunctive relationship for the Collider (see Equation B.9 in Appendix B). For the Order models this meant that the Collider likelihood was additionally distributed over patterns 5 and 7. For the Delay models this meant t_E was caused by the earlier-arriving of its two causes. Individually, 12 participants' judgments were closest to Order_N assuming a conjunctive Collider and 10 assuming a disjunctive Collider. Two participants were better fit by Order_S and seven by Delay_P.

Overall, there was relatively little sensitivity to the exact timing at which B occurred. If we compare patterns 2 to 6, we see that the chain was the modal response across early to late occurrence of B consistent with both Order_N and Order_S predictions. Notwithstanding the dominance of the Order_N model in explaining predictions, there was some evidence of sensitivity to event timings. Figure 6.8 shows participants' probability assignments to the Collider, chain, and fork for clips 1–7. For clip 2 where B happens

right after A, participants assigned some probability to the Collider structure. For clip 6 where B happens right before E, participants assigned probability to the fork. This timing sensitivity is revealed by fitting mixed-effect models to the points assigned to the Collider, the chain, and the fork across clips 2–6, with random means for participants. All three structures' assignments vary across these clips (Collider: $\chi^2(4) = 12.7, p < .013$; AB-chain: $\chi^2(4) = 9.5, p = .05$; A-fork: $\chi^2(4) = 27, p < .0001$) while the order models do not differentiate between these clips. Furthermore, we see hints of the bimodal shape for Collider assignments predicted by Delay_P. For the model this is a consequence of the conjunctive combination function (Equation 6.6) under which clip 5 is consistent with equal (e.g. 275ms) causal delays $t_{A\to E}$ and $t_{B\to E}$ with A's influence arriving earlier and "waiting" for B's, while not a perfect match to either chain or fork. As expected, chain judgments peaked when t_{AB} and t_{BE} are the same (clip 4) and the fork when t_{AE} and t_{BE} are the same (clip 7).

6.4.3 Discussion

In Experiment 6, we saw that participants' one-shot structure judgments were well explained by a simple model that only uses event order. The model predictions were not perfect though. Order_N underestimated participants' strength of preference for the Collider in clip 1, chain in 3-5 and fork in clip 7, and assigned more weight overall to the A- and B-singles. One possible explanation is that participants might have found some of the structures more or less likely a priori than others. Alternatively, participants might have expected A-single and B-single devices to generate clips in which one of the causal components never occurs even though they were told this would not happen in the instructions. Furthermore, the fact that A and B are perfectly simultaneous in clip 1, might have been seen as evidence for a common causal mechanism — for example some prior mechanism that ensures that the joint causes in the Collider occur in lock-step rather than occurring independently at different times.

The fact that participants' structure preferences were stronger than what was predicted by Order_N might relate to the fact that they replayed each clip several times. Some participants might have treated this as repeated evidence leading to stronger predictions. However, this does not explain the spread of probability in clips 2 and 6.

Participants' judgments shifted over clips 2 to 6 as predicted by Delay_P. This is evidence for some sensitivity to timing, however it was not sufficient to alter many participants' modal judgments away from those predicted by Order_N. Figure 6.8 shows an inverted

U pattern for the chain across clips 2 to 6, rather than the inverted V shaped curve predicted by Delay_P. An explanation for this is that people have limited ability to detect differences between interval lengths, with the modest differences between t_{AB} and t_{BE} in clips 3–5 falling below this threshold. Generally, participants exhibited a robust preference for the chain structure whenever activations occurred sequentially.

In Experiment 6, participants had very little evidence to go on. Having observed a device in action only once, one cannot experience its full range and variability in behaviour. Furthermore, single observations limit the scope for forming expectations about delays. In fact, the timings in Experiment 6 were only useful predicated on the Delay_P assumption that all cause–effect relationships between the components within a device have the same means and variances.[9] Thus, to better investigate the adequacy of the Order and Delay models, we now turn to extended learning, where participants observe multiple different clips of the same device and need to integrate the evidence to narrow in on the true causal structure.

6.5 Experiment 7: Integrating evidence

In this experiment participants saw several different clips for each causal device. To explore how participants integrated evidence, and to separate the predictions of our two order-based models Order_N and Order_S, we manipulated the order in which components activated during each clip. Participants saw several pieces of evidence, made an initial judgment, and then were able to update their judgment after some additional evidence. This procedure allows us to explore how learners revise their beliefs as they receive more evidence.

We hypothesised that participants' deviations from model predictions in Experiment 6 could be partly due to their having different assumptions about which structures are a priori more likely than others. Another possibility is that while many participants may be relying on temporal order, they might still distribute their likelihood differently than simply dividing it evenly across order-consistent patterns, in particular they might think qualitative patterns that imply reliable delays are more likely than those that do not. We test both of these questions directly in Experiment 7 by eliciting participants' priors and

[9] Although we note that participants could have formed delay expectations across the task in a hierarchical model fashion.

order-dependent likelihoods alongside having them make posterior judgments. This allows us to assess the relationship between prior beliefs, assumptions about the likelihood of different patterns, and posterior inferences on the level of individual participants.

6.5.1 Methods

Participants

Forty participants (19 female, $M_{age} = 30.8$, $SD_{age} = 7.4$) were recruited from Amazon Mechanical Turk as in the previous experiments. The task took 27.0 minutes ($SD = 16.6$) on average and participants were paid at a rate of \$6 an hour.

Stimuli and model predictions

In this experiment, we created evidence sets for 8 different "devices". For each device, participants were presented with four patterns of evidence (see Table 6.3). They were asked to provide a first judgment after they had seen the first three patterns of evidence, and were then given the chance to update their judgments after having seen the fourth pattern. We selected patterns such that, for five of the devices, our models predicted a strong shift in belief between the first and the second judgment, while for the other three, little or no shift was predicted.

For example, for device 4 (Table 6.3) participants first saw patterns 1, 2, and 5 ($AB \succ E$, $A \succ B \succ E$ and $B \succ A \succ E$) resulting in a strong prediction by both the Order$_N$ and Order$_S$ models that participants will favour the Collider. Finally, participants saw pattern 4 ($A \succ E \succ B$) which is incompatible with the (conjunctive) Collider model, meaning that both models predict a dramatic shift to A-single — the only remaining structure that is consistent with all four patterns (Figure 6.3 middle row). For three of these five devices the same shift was predicted by both Order$_N$ and Order$_S$, whereas for the other two a different shift was predicted. We only used sets of patterns that did not lead any of the considered models to rule out all the causal structures.

In addition to whether each set of patterns led to a large predicted shift between participants' first and second judgments, we also selected sets of evidence for which the most likely structure differed depending on whether or not participants made the assumption that causes and effects can occur simultaneously. Thus, Order$_N$ and Order$_S$ disagreed

TABLE 6.2: Experiment 7: Possible Temporal Order Patterns

Pattern	1	2	3	4	5	6	7
Order	$AB \succ E$	$A \succ B \succ E$	$A \succ BE$	$A \succ E \succ B$	$B \succ A \succ E$	$B \succ AE$	$B \succ E \succ A$

TABLE 6.3: Experiment 7: Evidence Sets (1^{st} - 4^{th} Piece of Evidence) for the 8 Devices

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1^{st}	1	2	2	1	2	2	2	2
2^{nd}	2	2	3	2	2	2	2	2
3^{rd}	5	2	4	5	2	2	2	2
4^{th}	2	2	3	4	4	5	1	6
Shift	N	N	N	Y	Y	Y	Y	Y
Different				N	N	N	Y	Y

Note: The numbers in the rows from 1^{st} to 4^{th} refer to the temporal order patterns shown in Table 6.2. The roles of components A and B were counterbalanced (e.g. pattern 2 $A \succ B \succ E$ becomes pattern 5 $B \succ A \succ E$) and responses re-coded. *Shift* shows whether a change of MAP judgment is predicted by one or both Order models (N)o/(Y)es. *Different* shows whether this shift is predicted to be different between Order_N and Order_S.

about the most likely structure for one or both judgments on 2 of the 8 devices (see Figure 6.3c for an example).

Since we elicit individuals' priors and order-based likelihoods, we can construct an individual order-based model Order_{IV} that makes predictions about $P_{IV}(M|\mathbf{d})$ given the qualitative order of events and each participant's subjective likelihoods $P_{IV}(\mathbf{d}|M)$ and prior $P_{IV}(M)$.

In the experiment, we drew intervals between components independently, effectively averaging out any effect of specific timings at the group level. Because each participant experienced different timings, and might have different priors, the Delay_P model also makes slightly different predictions for each participant.

Procedure

After reading the instructions, participants had to successfully answer comprehension check questions to proceed. The order in which the devices were presented was randomised between participants. However, the order of clips for each device was always as shown in Table 6.3. We varied the interval between each activation, drawing each from a uniform distribution between 200 and 1200 ms. The clips used in the experiment varied in total length between 1189 and 3094 ms depending on these intervals and whether there were three staggered component activation events (patterns 2, 4, 5 and 7) or only

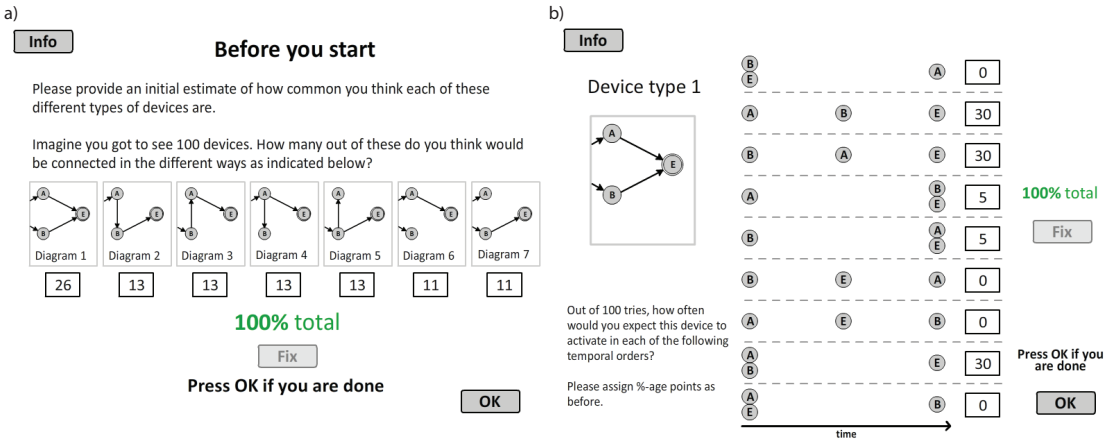


FIGURE 6.9: Experiment 7 interface. a) Eliciting priors before main task b) Eliciting likelihoods after main task.

two (patterns 1, 3 and 6). We counterbalanced two presentation orders of the seven structure hypotheses shown at the top of the screen between participants (Figure 6.6a).

In addition to the *posterior judgment phase* from Experiment 6, we added an initial *prior judgment phase* in which participants were asked to assign 100% points across the seven structures to indicate how probable they thought each of the different structures was a priori (see Figure 6.9a). In the *posterior judgment phase*, participants made judgments for 8 devices. They were provided with the qualitative visual summary of the clips they had seen. Finally, participants completed a *likelihood judgment phase*. In this phase, participants made seven additional percentage allocations, one for each causal structure. For each allocation, they were shown one of the seven structure diagrams. They were then asked: "Out of 100 tries, how often would you expect this device to activate in each of the following temporal orders?" Participants distributed 100%-points across the different temporal order patterns (see Figure 6.9b). The order in which participants were asked about each structure, and the order in which the different temporal patterns appeared on each page were randomised between subjects.

When making their posterior judgments, participants were provided with a qualitative summary of the clips they had seen so far (similar to the those in Figure 6.12a).

Participants were instructed that clicking on the "Start" button constituted the cause of any parentless components in the model. This was indicated in the structure hypotheses by the addition of arrows connecting to any parentless components in each diagram (cf. Figure 6.9a).

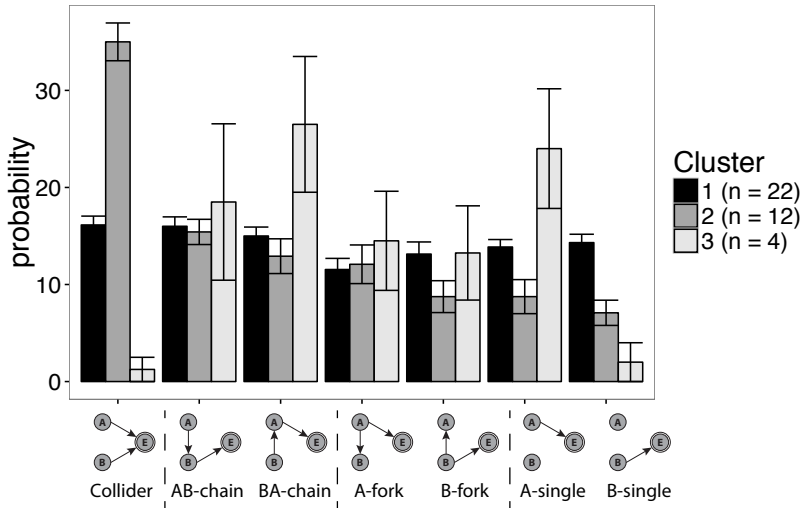


FIGURE 6.10: Elicited priors split into clusters as detailed in text. Error bars show standard errors.

6.5.2 Results

We will discuss the results from the *prior judgment phase*, *likelihood judgment phase*, and *posterior judgment phase* in turn.

Prior judgment phase

21 of the forty participants' priors differed significantly from a uniform according to χ^2 tests, with Bonferroni corrected significance (i.e. $p < \frac{.05}{40}$). After removing two participants who assigned 0% to more than half of the structures, we performed a cluster analysis on the remaining 38 participants, finding three clusters.[10] Twenty-two participants assigned roughly equal weight to all seven options (see Figure 6.10). Twelve assigned approximately double to the Collider compared to the rest of the structures. Four other participants formed a third cluster with no apparent systematicity in their priors.

The 12 participants who gave more mass to the Collider structure might have been thinking in terms of *types* of structure, dividing evenly across Colliders, chains, forks and single, then subdividing within each type. This could explain their putting more prior weight on the Collider, since it is the only structure within its class. By splitting the resulting probabilities across class members, the Collider ends up with greater prior probability due to being a unique member of its class.

[10] This was established by fitting a Gaussian finite mixture model using R's `mclust` package.

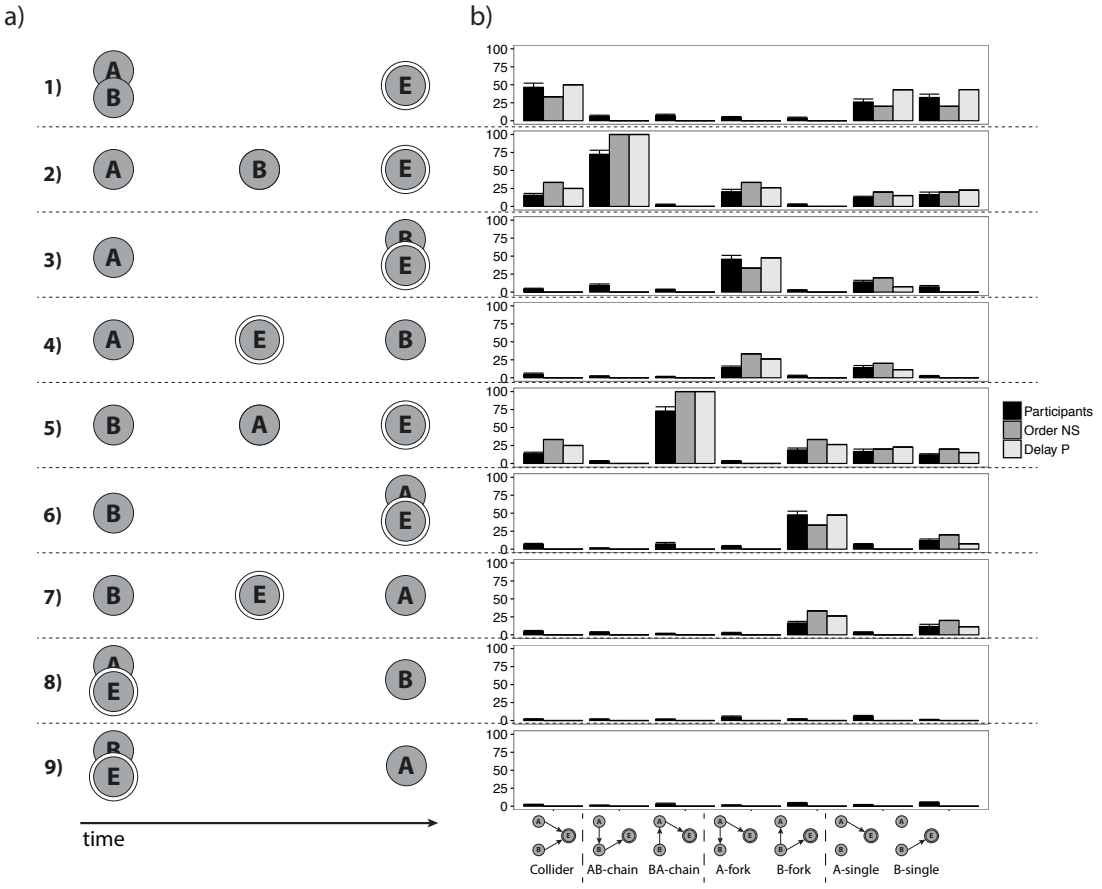


FIGURE 6.11: Elicited likelihoods. a) Nine temporal order patterns b) Participants' elicited likelihoods compared with those of Order_N and a variant of Order_N that distributes likelihoods across *types* of consistent patterns — lumping together $A \succ B \succ E$ and $B \succ A \succ E$ for the Collider, and $A \succ B \succ E$ and $A \succ E \succ B$ for the forks. Error bars show standard errors.

Likelihood judgment phase

TABLE 6.4: Experiment 7: Likelihood Judgment Model Fits

Model	r	r_s	Mode match	RMSE	N
Baseline			11%	15.4	1
Order_N	0.92	0.78	71%	8.9	11
Order_S	0.57	0.80	43%	12.8	4
Delay_P	0.98	0.81	100%	7.3	24

Note: r = average Pearson's r correlation between average assignments to structures within each device and model predictions. r_s = average Spearman's rank correlation within problems. Mode match = proportion of problems where participants' modal choice matched model's. RMSE = root mean squared error. N = Number of individuals best correlated by model.

Likelihood judgments were most highly correlated with marginal likelihoods of the patterns under Delay_P ($r = .98$), followed by Order_N with Order_S considerably lower (see

Table 6.4).[11] Inspecting Figure 6.11, reveals that the Delay_P based likelihoods captured the fact that participants assign more probability to the patterns implying reliable delays (more to pattern 1 than patterns 2 or 5 for the Collider, and more to pattern 3 than 2 or 4 for the A-fork, and similarly for the B fork).

To check whether participants largely made the same assumptions about the devices as our models, we checked how frequently they assigned likelihoods to patterns ruled out under all of the models we consider. Overall, participants assigned much less likelihood to these patterns (10.5%) compared to the 44% expected from random allocation. However, eighteen participants assigned some likelihood to patterns ruled out by Order_N, Order_S and Delay_P, assigning an average of $7.4 \pm 13\%$ of their points to 5.3 ± 10 of the 24 patterns. Of these, the most frequently were $A \succ E \succ B$ and $B \succ E \succ A$ under the Collider, with nonzero likelihoods assigned by 12 and 14 participants respectively. As a result there was a higher probability of assigning non-zero likelihoods to patterns ruled out by our models under the conjunctive Collider than on average over the other structures $\chi^2(1) = 5.1, p = 0.02$. This confirms our suspicion that some participants did not make the conjunctive assumption when reasoning about the Collider, in spite of instructions.

Posterior judgment phase

We analysed the posterior judgments by comparing linear mixed models with random intercepts for participants, and structures within participants. By design, neither device (1:8) nor judgment (1^{st} vs. 2^{nd}) can have a main effect on assignment of % points. This is because judgments were constrained to add up to 100% across the structures. Instead, effects are indicated by interactions between these different factors and the assignments across the structures. Structure interacted with device $\chi^2(55) = 1384, p < .0001$, confirming that judgments were affected by the different evidence sets. Judgment (1^{st} versus 2^{nd}) also interacted with device $\chi^2(7) = 99, p < .0001$, and there was a three-way interaction between judgment, structure and device $\chi^2(49) = 286, p < .0001$ confirming that the impact of the final piece of evidence was different for some devices than others. The complexity of these interactions prohibits direct interpretation but we can compare judgments' to the predictions of Bayesian updating based on participants' elicited priors and likelihoods, either with or without additional sensitivity to the intervals.

[11] We assumed the same parametrisation as in Experiment 6, and encoded the timings implied by the depictions of the order patterns (e.g. Figure 6.11a) assuming they represented a total interval of 1400 ms, with 700 ms between the initial and middle events for patterns 2, 4, 5 and 7, corresponding to the mean interval between events in the task.

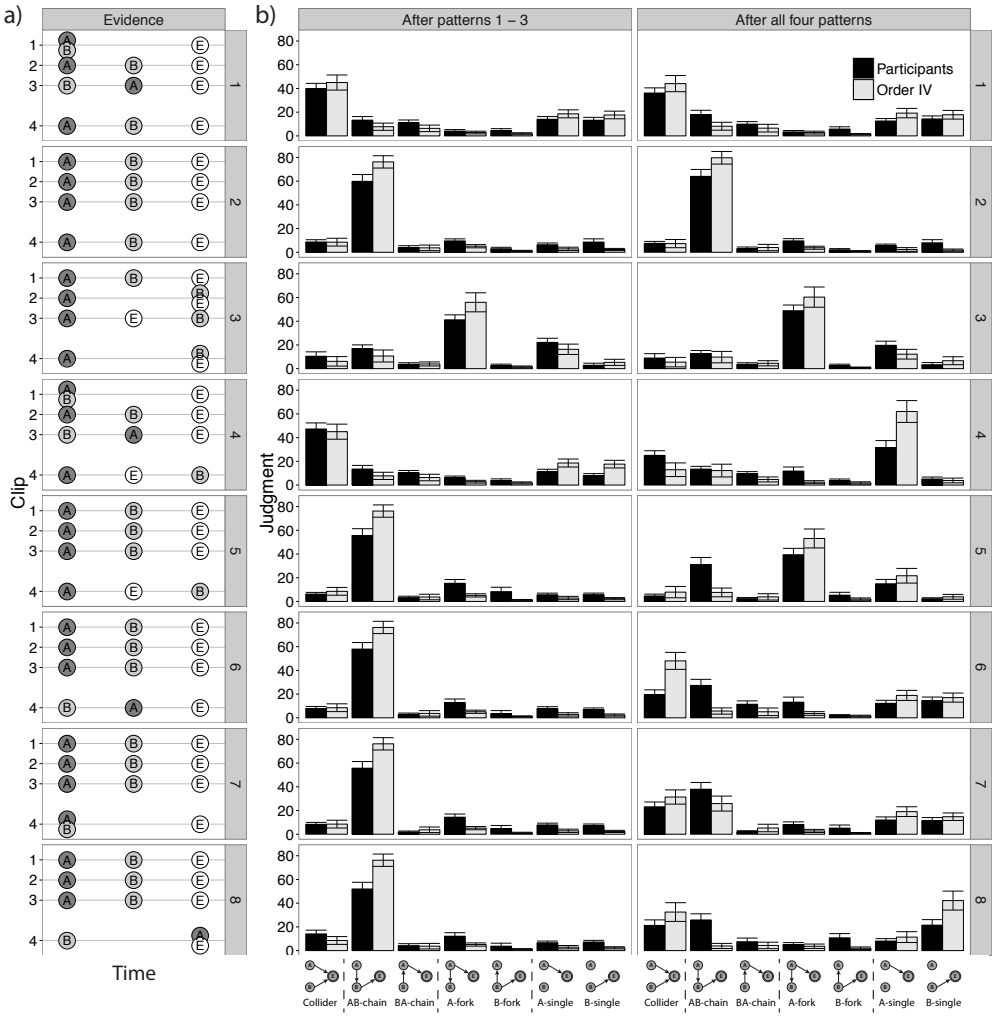


FIGURE 6.12: Experiment 7 posterior judgments. a) Devices and qualitative order of activations for each. *Note*: Exact timings were drawn at random from Uniform(200, 1200) for each participant in this experiment and so are not shown in full. b) Participants' posterior judgments (black bars) compared to a model based on individually elicited priors and order-based likelihoods Order_{IV} (grey bars). Left hand column, judgments after viewing 3 clips, right hand column judgments after all four clips. The Order_{IV} bars omit cases in which participants' chosen likelihoods and priors led to all hypotheses being ruled out. Error bars show standard errors.

Participants' average posteriors were very closely correlated with the predicted average over posteriors based on the priors and order-based likelihoods they provided (Order_{IV}). By computing these posteriors then averaging over participants, we get a $r = .95$ correlation with judgments and a RMSE of 7.0% compared to baseline of 14.3%. It does not make sense to average the Delay model posteriors in this experiment since timings differed between participants. However, we can check for timing sensitivity at the level of individuals. Here, we find that most participants' posteriors are still best described by the Order_{IV} model that combines the priors and order-based likelihoods they provided

themselves (28/40).[12] However 10 were better described by the posteriors under Delay$_P$ suggesting some additional sensitivity to experienced timing.

Inspecting Figure 6.12 we see that aggregated participant posterior judgments are typically a little less peaked than the aggregated Bayesian posterior predictions, even though these were based on their priors and likelihoods, and account for the heterogeneity of assumptions people made about the task. In particular, where we chose a fourth clip such that we predicted a modal shift between the first and second judgments (devices 5-8), we see considerable residual percentage points for the previously favoured structure to that which the models favour after. For example, for device 5, participants' priors and likelihoods suggest they should strongly favour the A-fork after viewing the final clip, but participants move only around half the probability mass, leaving a considerable amount "behind" on the previously favoured AB-chain, which all our models consider to be ruled out. This suggests that participants were generally somewhat conservative in their updates. Their beliefs were moved less by the evidence than their priors and likelihoods would suggest they should be (Edwards, 1968). To test this more thoroughly we considered a variant of Order$_{IV}$ that updates its beliefs conservatively.

We can model conservatism within the Bayesian framework through addition of unbiased noise to participants' likelihood functions, such that patterns that should be ruled out by a structure given one's assumptions, instead retain some ϵ probability. If participants are generally conservative, we expect such a model that incorporates noise into the likelihoods for each observation to better explain their final judgments.

To do this, we created noisy likelihoods by mixing each participant's reported likelihood function with uniform likelihoods (with $\frac{1}{9}$ over the 9 patterns of data participants distributed over for each structure) to a degree controlled by a free parameter $\epsilon \in [0, 1]$ (i.e $P_{IV}(\mathbf{d}|s)^{cons} = (1-\epsilon)P_{IV}(\mathbf{d}|s) + \frac{1}{9}\epsilon$).[13] We fit ϵ to each participant's data separately, by maximising the correlation between the prediction given Bayesian integration of the priors and likelihoods they reported, and their own posterior judgments.[14] We found that 32/40 participants had a non-negligible best-fitting ϵ parameter (> 0.01), indicating conservatism in their evidence integration relative to the Bayesian ideal. The mean ϵ

[12] For 14% of first and 27% of second judgments, all structure hypotheses were ruled out based on combining individuals' priors and likelihoods. This happened at least once for 19 out of the 40 participants. To allow comparison we simply had the Order model predict a uniform distribution over the hypotheses in these cases, guaranteeing a correlation of 0 for that device — the same as the Baseline model.

[13] For $\epsilon = 1$, $P_{IV}^{cons}(\mathbf{d}|M)$ is uniform and therefore results in no belief change.

[14] We use the Brent (2013) algorithm to do the optimisation as implemented by R's `optim` function.

was .36 ($SD = .35$). Inclusion of conservatism increased the aggregate model correlation of Order_{IV}^{cons} to $r = .97$, RMSE= 6.5% compared to Order_{IV}'s $r = .95$, RMSE= 7.0%.

6.5.3 Discussion

In Experiment 7, we attained a clearer picture of the sources of variability in people's causal structure inferences. Many participants reported priors that distributed probability mass uniformly at the level of *types* of structures rather than response options. The Collider was the only unique structure (since there were two chains, two forks and two singles), and it was judged to be a priori more likely than the rest of the structures by many participants. This suggests that these participants generated uniform prior probabilities based on more abstract representations of the causal structures and evidence patterns, rather than taking the option set we provided as distributionally representative.

Participants found structures that exhibited equal delays more likely than unequal delays. We were able to capture this very well by our Delay_P model which favours structures that imply causal delays that are more similar on average (with a $r = .98$ correlation with the aggregate patterns and a better fit than the other models we considered for 24/40 participants). Participants' made these likelihood judgments after having completed the posterior judgment phase. It is thus possible that they tried to make their likelihood judgments consistent with the posterior judgments they had provided in the previous phase of the experiment.

Interestingly, despite distributing likelihoods in a way that suggested they preferred equal delays across devices' components, participants still appeared quite insensitive to exact event timings. The majority of participants' posterior judgments were better described by Order_{IV} than Delay_P suggesting that participants paid little attention to exactly how far apart in time the events were in the clips. We note here though that the design of the experiment might have nudged people toward this behaviour. We provided summaries showing the qualitative order of events in Experiment 7 while the exact event timings were only represented in the clips themselves. This may have encouraged participants to focus predominantly on order. Furthermore, by selecting clips that provided lots of order information, the resulting data was not distributionally representative of reliable generative gamma delays.

We found that we can capture participants judgments even better by positing that they were somewhat conservative in their integration of the evidence they observed,

over and above what was implied by the likelihoods they provided. Conservatism relative to Bayesian predictions is a consistent psychological finding (Bramley, Lagnado, & Speekenbrink, 2015; Edwards, 1968; Fischhoff & Beyth-Marom, 1983). In this task, it could reflect a number of things. Participants may have suspected that the devices might change structure over time, and so not want to rule out a possibility that could later be true. They might also distrust what they were told in the instructions, have forgotten or be unsure about them (Corner, Harris, & Hahn, 2010). Under-updating of judgments might more fundamentally be a consequence of their processing limitations, either directly, or as a way of compensating for the possibility of having made perceptual or memory errors about the evidence they had seen.

Our qualitative order models did well in explaining participants' inferences in the tasks we have looked at so far, even explaining evidence integration over multiple trials where there is, in principle, enough timing evidence to start to form expectations about the delays. However the experiment emphasised order information by using non-representative delays and providing qualitative visual summaries of the evidence during posterior judgments, yet Delay_P still outperformed Order_N for some participants. Finally, the close correspondence between Delay_P and participants' qualitative likelihood ratings clearly show that timing matters, even if their role here was predominantly limited to shaping peoples' order expectations.

To look more closely at the role of timing, we now turn our focus to a situation where order is non-diagnostic and the only available information comes from the variability and correlation in event timing. This will allow us to assess the extent to which people are capable of using timing information at all, and the adequacy of our normative model in capturing the ways in which people use temporal information.

6.6 Experiment 8: Learning from timing variability alone

In this experiment, we focus on causal inference from timing variability alone. To isolate timing from order cues, we chose a more constrained situation than before, with only two possible structures (an $S \rightarrow A \rightarrow B$ chain and an $A \leftarrow S \rightarrow B$ fork) and evidence where the order of activation of three components was (almost) always the same ($S \succ A \succ B$). We systematically varied the mean and variability of the inter-event timings such that they were more consistent with having been generated by either a chain or a fork under the Delay_I assumption as we describe below.

We hypothesised that participants would be sensitive to these differences and able to use them to distinguish between the two candidate structures. However, we also expected based on the results from the previous experiments, that participants would have a general preference for the chain. While the chain can only produce the $S \succ A \succ B$ pattern, the fork is more flexible. We also hypothesised that participants would find it more difficult to draw inferences from quantitative differences in time intervals, versus the more obvious and definitive qualitative differences in event order. Thus, we predicted that participants would be more uncertain overall in their posterior judgments. To assess how well participants detect and track timing variability across tests and hypotheses, we first elicited judgments based on simply experiencing the timings. Afterwards, we provided participants with summaries of the trials detailing all the timings visually, and allowed them to update their judgments. The idea was that providing participants with summaries would eliminate any potential memory effects, or effects resulting from perceptual noise associated with encoding the timings, providing a helpful comparison to the judgments based on experience alone. Generally, we expected participants' preference for one of the two structures to become stronger and closer to normativity after having seen the summary.

A further question is whether participants who are able to learn the true causal model are also able to learn the causal delays, such that they can make predictive judgments about what patterns of evidence the device is likely to produce in future tests. To explore this question, the experiment included an additional task where participants had to make a predictive judgment.

6.6.1 Methods

Participants and materials

104 University College London undergraduates (87 female, $M_{age} = 18.8$, $SD_{age} = 0.81$) took part in this experiment under laboratory conditions as part of a course requirement. The task took 23.0 minutes ($SD = 3.1$).

Stimuli

Participants had to judge whether a device was a $S \rightarrow A \rightarrow B$ chain or a $A \leftarrow S \rightarrow B$ fork. Both chain and fork structures shared an $S \rightarrow A$ connection, but differed in

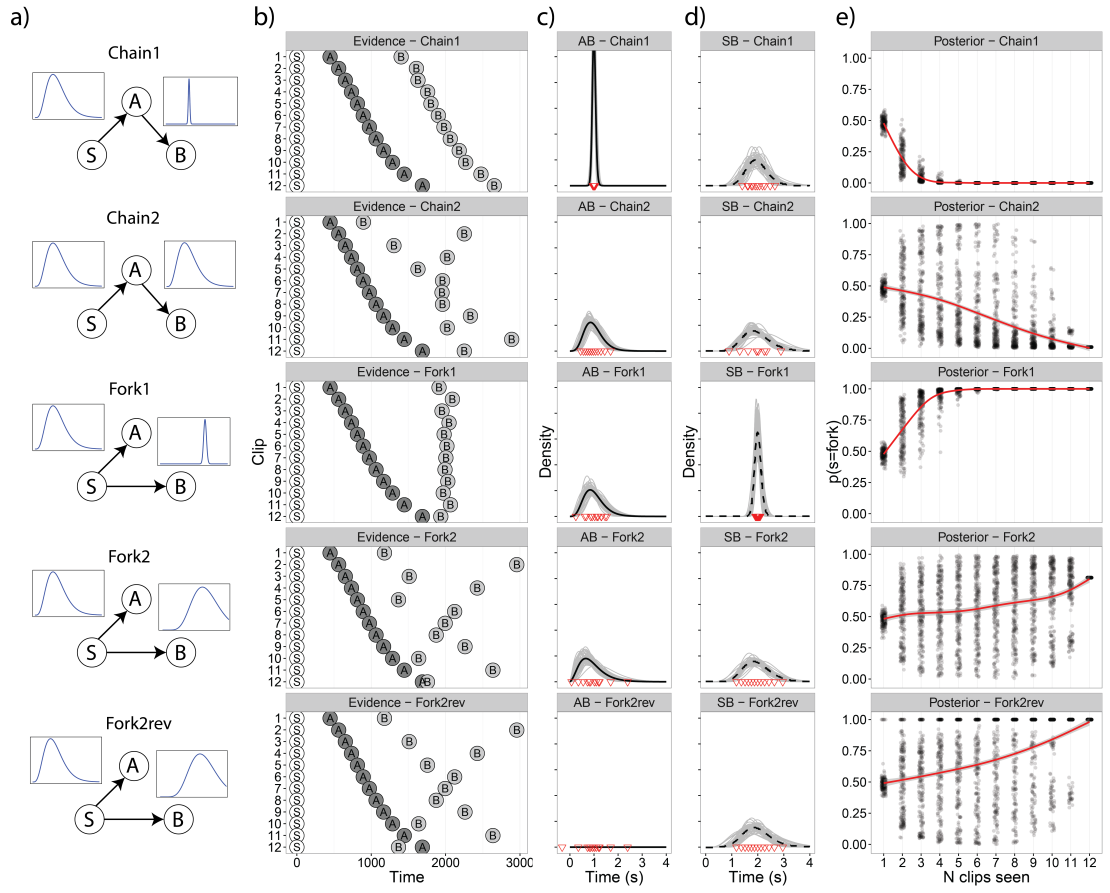


FIGURE 6.13: Experiment 8 stimuli and model predictions. a) Graphical representation of the five device types. b) Plot showing the 12 patterns generated for each device. c) Red inverted triangles: t_{AB} for patterns 1:12. Gray lines: $P(G_{A\to B}|\mathbf{d})$ for a posterior sample of αs and βs. Dashed black line: The posterior marginal likelihood of $G_{A\to B}$. d) As in c) but for $G_{S\to B}$ under the fork structure. e) Posteriors $P(m = \text{Fork}|\mathbf{d})$ for progressively more evidence. Individual dots for the samples of evidence seen by participants, lines smoothed average (using the general linear additive model with integrated smoothness estimation `gam` from R's `mgcv` library). *Note*: Individual points are jittered to increase visibility.

whether they had an $S \to B$ or an $A \to B$ connection. This implies that t_B could be explained by one of two delay distributions: either $t_{S\to B}$ or $t_{A\to B}$. Under the independent Delay$_I$ model, this results in a preference for one of the two structures, depending on which of these inferred delay distributions can assign more likelihood to the evidence (marginalising over its unknown parameters).

In order to construct the evidence, we first created two generative chain (Chain1 and Chain2) and fork devices (Fork1 and Fork2) by augmenting each connection with a delay distribution (see Figure 6.13a). All four devices shared an $S \to A$ connection with delay distribution $G_{S\to A}(\alpha = 5, \mu = 1000\text{ms})$. Concretely, this meant that A would occur an average of 1000ms seconds after S but with considerable variability. We then chose distributions for $A \to B$ for the chains and $S \to B$ for the forks such that the interval

between S and B t_{SB} was 2000ms on average, but the shape and extent of the variability in the timing of B depended on the underlying connections.

In Chain1 there was a near-constant $G_{A\to B}(\alpha = 1000, \mu = 1000\text{ms})$, while in Chain2 $t_{A\to B}$ had as much variability as $t_{S\to A}$. In Fork1, the $S \to B$ connection had a near-constant 2 second delay $G_{S\to B}(\alpha = 1000, \mu = 2000\text{ms})$ while in Fork2 the delay was variable $G_{S\to B}(\alpha = 10, \mu = 2000\text{ms})$.

We used these four generative devices to select sets of 12 clips used as evidence. To ensure that the selection of clips was representative for the generating distributions, we took 12 equally spaced quantiles from each distribution.

To ensure that the delay draws for $G_{S\to B}$ (or $G_{A\to B}$ for the forks) were independent of those for $G_{S\to A}$, they were paired in counterbalanced order. The resulting sets of evidence are depicted in ascending order of t_{SA} in Figure 6.13b. Finally, we included a variant of Fork2, named Fork2rev, which included a single order reversal trial. This allows us to compare the respective strengths of order and timing cues.

Model predictions

We used Delay_I to obtain a posterior joint distribution over the true structure (i.e. fork or chain) and its associated parameters.[15] We obtained posterior predictions by averaging over the parameters. These predictions are normative in the sense that the Delay_I model inverts the true generative model. Figure 6.13e shows how these predictions change with each additional clip seen. Because we randomised the order of the clips, there is variability in what evidence the model has received so far. Each point in the plots shows the predicted posterior given the evidence an individual participant has seen up to this point. The red line shows the averaged predicted posterior. By the 12th clip, all participants have seen the same evidence so the predictions converge.[16]

Figure 6.13e shows that the model rapidly infers that the true model is a chain for Chain1 and a fork for Fork1. Looking at the predictive distribution subplots (Figure 6.13c and d), we see that this is due to the model's ability to fit a tighter distribution onto the experienced timings under the true model, assigning less mass to all the data points while they are more spread out and unevenly distributed under the alternative structure.

[15] We used the Delay_I variant of our delay model because the Delay_P variant assumes that all delays share the same parameters, and participants were explicitly instructed that this was not the case.

[16] We used MCMC to estimate these posteriors without specifying any prior on delay parameters. In the appendix we compare these to Simple Monte Carlo sampling predictions under a variety of priors. This allows us to assess the impact of prior choice in Experiments 6 to 8.

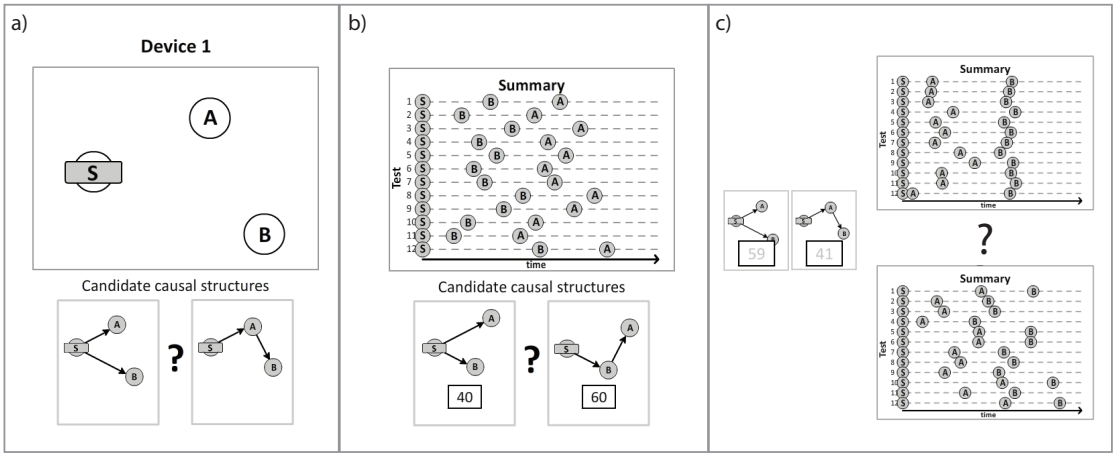


FIGURE 6.14: Experiment 8 interface. a) Testing the device b) Viewing a visual summary c) Making a predictive judgment.

Under the noisier Chain2 and Fork2 evidence, the model forms the correct preference but does so much more slowly, retaining significant uncertainty even after 12 clips for Fork2, where the delay distribution is only slightly less variable under the fork structure than the chain. Finally, for Fork2rev the predictions are the same as Fork2 until the order reversal trial is seen and the chain is ruled out. This becomes increasingly likely on later trials and certain after all 12 clips. Thus, normatively we expect more points to be assigned to the chain structure for Chain1 and Chain2, than for Fork1, Fork2 and Fork2rev; more to Chain1 than the more difficult to infer Chain2. Likewise, we expect more points to be given to the fork structure for Fork1 than Fork2. Finally, since the order cue in Fork2rev rules out the chain we expect judgments here to be more strongly in favour of the fork structure than for the other fork patterns.

Procedure

Participants were instructed about the two possible causal models, the interface, the number of problems they would face, the number of tests they would perform for each problem, the presence of delay variability, and the independence of variability between different connections. Participants initiated the system by clicking on the "S" component and watching when the other two components activated (see Figure 6.14a). To familiarise participants with the delay variability, they interacted with four two-component devices during the instructions, each with a single cause and a single effect. They tested each device at least 4 times. There were two pairs with short ($\mu = 1s$) delays, one near-constant and one variable, and two with longer ($\mu = 2s$) delays, likewise one near constant and one variable. Participants were also instructed that the variability of the

delays of the different components of a device were independent such that an unusually long $t_{S\to A}$ would not imply that there would be an unusually long $t_{S\to B}$ or $t_{A\to B}$. Before proceeding to the main task, participants had to correctly answer comprehension check questions.

All participants faced each of the 5 problem types twice, once as detailed in Figure 6.13 and once with the labels and locations of A and B reversed (as in Figure 6.14b). Thus, there were 10 within-subjects test problems overall. On each test problem, participants watched 12 clips in a random order. For each problem they made 3 causal judgments. They made their first judgment after the 6th clip, their second after all 12 clips, and a final judgment after seeing a visual summary of the timelines of the clips they had seen (similar to the quantitative summary in Experiment 7, see Figure 6.14b). Participants gave their causal judgments by distributing 100% points across the two structures. During trials 7–12, participants' initial response remained visible but greyed out in the response boxes. They then had to interact with one of the response boxes (changing the value or just pushing enter) to unlock the "Continue" button on the second and third judgments.

In addition to eliciting structure judgments for 10 problems within subjects, we also elicited predictive judgments on one additional problem which was varied between subjects. On this final problem, participants either saw evidence from Chain1 or Fork1, in a new order. We selected which evidence was seen at random between subjects (45 out of 104 subjects saw Chain1, the rest saw Fork1). The first and second judgments were identical to the previous problems, but instead of seeing the visual summary, participants were presented with two side-by-side visual summaries of new draws of 12 patterns, one generated by a Chain1 structure and one by a Fork1 structure (See Figure 6.14c). They were then asked to distribute 100% between the two sets of evidence indicating which evidence was more likely to be produced by the current device.

6.6.2 Results

Structure judgments

In order to analyse participants' judgments, we must account for the fact that each participant faced each device twice. To do this, we fit linear mixed effects models to all judgments, with participant and device as random effects. To test our specific hypotheses about the differences between devices, we constructed four orthogonal contrast codes.

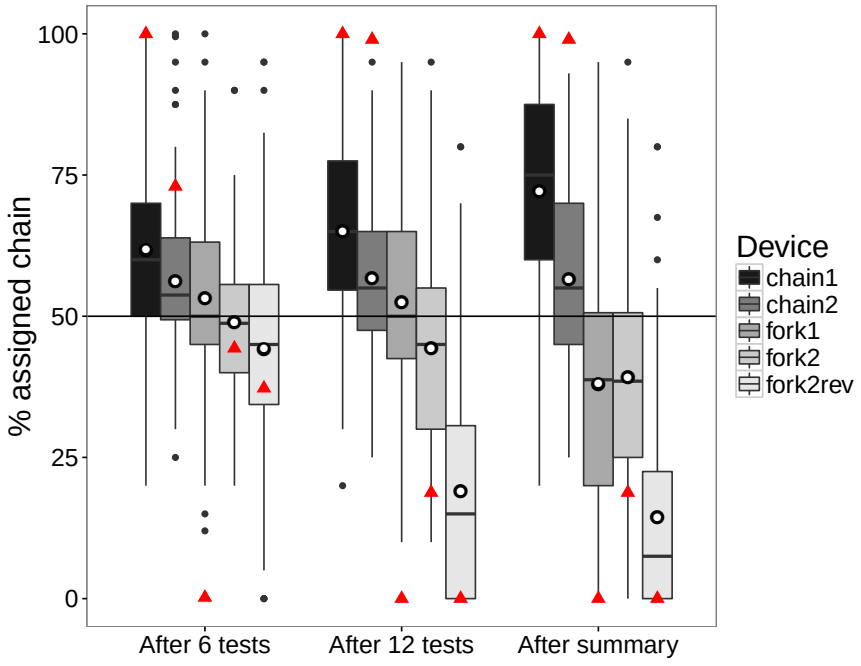


FIGURE 6.15: Judgments for the different devices. Boxplots show participants' median and upper and lower quartiles, participants with judgments ± > 1.5 interquartile range are plotted separately. White filled black circles = participant means. Red triangles = Delay_I posteriors.

TABLE 6.5: Experiment 8: Main Effects and Planned Comparisons for First, Second and Third Responses

	Response 1	Response 2	Response 3
Main effect (LR)	86***	334***	378***
Planned contrasts			
Intercept	$52 \pm .9\%^{***}$	$46 \pm .9\%^{***}$	$42 \pm .9\%^{***}$
1. Chains vs. Forks	$10.2 \pm 1.3\%^{***}$	$22.2 \pm 1.3\%^{***}$	$33.8 \pm 1.6\%^{***}$
2. Chain1 vs Chain2	$2.7 \pm 1.0\%^{**}$	$4.1 \pm 1.1\%^{**}$	$7.8 \pm 1.3\%^{***}$
3. Fork1 vs Fork2	$2.1 \pm 1.0\%^{*}$	$4.0 \pm 1.1\%^{**}$	$-.6 \pm 1.3\%$
4. Forks1&2 vs Fork2rev	$4.5 \pm 1.1\%^{**}$	$19.6 \pm 1.3\%^{***}$	$16 \pm 1.5\%^{***}$

Note: For main effects we report the likelihood ratio for a model with device type as predictor relative to a model with just an intercept. For each planned comparison we report the size of the effect (%) ± standard error, and level of significance: $^{*} = p < .05$, $^{**} = p < .01$, $^{***} = p < .001$.

These compared: 1. [Chain1, Chain2] to [Fork1, Fork2 and Fork2rev], 2. Chain1 to Chain2, 3. Fork1 to Fork2 and 4. [Fork1, Fork2] to Fork2rev, matching the predictions described above. The four regressions are summarised in Table 6.5. All three judgments differed by device type, with the size of these differences increasing for the judgments made after performing 12 compared to 6 tests, and after seeing the visual summary relative to before. For instance, participants assigned 10.2% more percentage points to the chain diagram when the true structure was a chain, after 6 tests, increasing to 22.2% after 12 tests and to 33.8% after viewing a visual summary of the evidence. On all

three judgments, participants assigned significantly more points to the chain diagram for chains than forks. They also assigned more chain points to the reliable than the unreliable chain, and more to the forks that did not exhibit the order cue (Fork1 and Fork2) compared to Fork2rev. However, on judgments after 6 and 12 tests, participants assigned *more* points to the chain diagram (i.e. fewer to the fork) for the theoretically easier and "reliable" Fork1 than the "unreliable" Fork2. After the visual summary, the fork diagram was equally favoured for each of these two devices.

Looking closely at the evidence we generated, we see that the difference between Chain2 and Fork2 is very subtle. While the t_{AB} interval is more variable under Fork2 than Chain2 (Figure 6.13c 2^{nd} vs 4^{th} row), t_{SB} is actually also slightly more variable under Fork2 than Chain2 (Figure 6.13d). Thus, if participants focused only on t_{SB} we would expect them to favour the chain structure for this problem. The fact that participants still form a preference for the chain for Chain2 and the fork for Fork2 based on this subtle difference in t_{AB}, while failing to note the reliable t_{SB} in Fork1, is suggestive that participants were particularly tuned to monitoring the successive intervals rather than the overall interval. We examine this idea in more detail in the General Discussion in the Section *Sensitivity to timing: Toward a process model.*

Predictions

For the final problem, participants had to predict which of two evidence sets was more likely to be generated by the device they had just learned about. Here, participants favoured the chain evidence marginally more when the true structure was Chain1 compared to when it was Fork1 $t(102) = 1.7, p = .044$ (one-tailed). Participants assigned significantly more than 50% to the chain evidence when the true structure was the chain $t(44) = 1.9, p = .029$ (one-tailed) but were not significantly more likely to favour the fork evidence for the fork device $t(58) = -.26, p = .36$ (one-tailed). However, participants' strength of judgment toward the chain (/fork) was not statistically related to their preference for the evidence actually generated by the chain (/fork) $F(1, 102) = .6, r = .08, p = .4$. For example, in Figure 6.14c the fact that this particular participant assigned 59% to the fork does not mean they will assign more predictive probability to the future fork-generated evidence (top) over the chain-generated evidence (bottom). We note, though, that we did not test participants' predictive knowledge very thoroughly in this experiment. Only a single predictive trial was included, varied between subjects, and there was no incentive or instruction for participants that they

should try to learn to predict the devices. Further research is needed to gauge the extent to which people can learn to predict the temporal dynamics of causal systems.

In sum, we found that people were able to distinguish between direct and indirect causation (i.e. a fork and a chain) based on the variability and correlation in event timings alone. However, people found this inference much tougher than making judgments based on having observed different temporal orders of events. In this experiment, some participants reported relatively weak preferences despite having seen considerably more data, and having fewer structure hypotheses to evaluate than in Experiments 6 and 7.

6.6.3 Discussion: Sensitivity to timing

Our Bayesian Delay_I model broadly captured aggregate judgments (see Figure 6.15). However, there is some evidence that participants may have solved the task in a more heuristic way. Firstly, participants' judgments were much less strong than the normative model's preferences. Secondly, participants had trouble predicting future evidence, suggesting they did not finish each problem with clear expectations about the device's delays. Third, Delay_I strongly favoured the fork structure after seeing only a few clips from Fork1, while participants remained at chance for this problem until the summary.

Ideal probabilistic structure inference involves maintaining a probability distribution over all candidate hypotheses. This is infeasible in the general case as there is a near-infinite number of possible models. There have been several recent proposals that people maintain a single candidate causal model at a time, stochastically switching when their current model proves strongly incompatible with evidence (Bonawitz et al., 2014; Bramley, Dayan, et al., 2017). Additionally, Lagnado and Sloman (2006) propose that people often take event order as an initial proxy for causal order. In this section we consider several heuristics based on the idea that participants in Experiment 8 used simpler statistics to identify the generative model without computing the predictions under both structures at once.

Does A predict B?

In general, if A causes B, we expect that the time at which we observe A (relative to its cause S) to be predictive for when we will later observe B (also relative to S). Thus, a reasonable proxy for computing the full posterior is to try and estimate the strength of this predictive signal. In the current context this comes down to a correlation between

t_{SA} and t_{SB}, hereafter $\text{cor}(t_{SA}, t_{SB})$. If $\text{cor}(t_{SA}, t_{SB})$ is positive, this is a sign that S's causing of B may be mediated via A — that is, observing an unusually early/late A is a noisy predictor of an early/late B (see Figure 6.13a). Conversely, if t_{SA} is statistically independent of t_{AB} this is more consistent with the idea that B is caused directly by S as in a fork structure.

Variance under a single structure

Computing a correlation between t_{SA} and t_{SB} across clips might still make too strong demands on perception and storage to be estimated online while watching the clips. The issue here is that the correlation depends on encoding two overlapping intervals for each test, storing them, and comparing their relationship across multiple trials. It is well-established that there are strong limitations on explicit attention and short-term memory which may prohibit such explicit multitasking (Baddeley, 1992; Lavie, 2005). Rather, it seems plausible that learners might only monitor the timings in the clips under a single hypothesis at a time, for example either focusing on t_{AB} if they are currently entertaining the *chain* structure, or t_{SB} if currently entertaining the *fork* structure.

Accordingly, a simpler strategy than comparing models would be to monitor the variance assuming that one or the other structure is true. If this variance seems "too high" one can reject the structure hypothesis and start monitoring the delays under the alternative structure.

Assuming that participants tend to perceive event order as causal order by default (Lagnado & Sloman, 2006), it is possible that participants found it more natural to monitor $\sigma^2(t_{SA})$ and then $\sigma^2(t_{AB})$ than to monitor $\sigma^2(t_{SB})$ (while ignoring the intervening event at A). Thus, $\sigma^2(t_{SB})$ may effectively have been masked by participants' default tendency to perceive succeeding events as a chain, and thus only encode the delays between directly succeeding events.

Online approximation

Estimating variance of the delays across trials may already be challenging. As we mentioned in the introduction, many models of sequential estimation avoid storing all the data, replacing an operation over all the evidence with a simpler adjustment that can be performed as evidence comes in (Halford, Wilson, & Phillips, 1998; Hogarth & Einhorn, 1992; Petrov & Anderson, 2005). We propose a simple model based on this idea here.

Average pairwise difference (APD) simply stores the difference between the interval in the latest clip t_{XY}^{k} and the one before t_{XY}^{k-1}, summing this up across trials. When variance is high this will tend to be high too but it is also sensitive to the order in which evidence is observed, being larger when intervals fluctuate more between adjacent tests.

Each of these measures — $cor(t_{SA}, t_{SB}), \sigma(t_{SA}), \sigma(t_{SB}), \sigma(t_{AB}), \text{APD}(t_{SA}), \text{APD}(t_{SB})$, and $\text{APD}(t_{AB})$ — assigns a value to the evidence seen at each time point by each participant. Thus, all the measures make different predictions for each participant on the first judgment because the clips seen so far differ between participants. Additionally, the APD measure is computed sequentially and thus creates order effects and results in different predictions for different participants for the second and third judgments, too.

We used all these measures as predictors of the number of percentage points assigned to the chain structure on each judgment with a prediction of zero indicating 50% chain (50% fork). This means that measures which support the chain have positive weights and measures that support the fork have negative weights, and the intercept indicates a baseline preference for one or the other model.

We hypothesd that one or a combination of these simpler measures $\sigma(t_{XY})$ or $\text{APD}(t_{XY})$ would capture participants' judgments better than the Delay_I posterior. Furthermore, we predict that most participants would base their judgments on the variance of t_{AB} rather than t_{SB}, given that it is easier to estimate the interval between subsequent events, rather than separated events. After the summary, we hypothesised that participants' judgments would become more normative, that is, closer to the predictions of Delay_I.

Modelling all participants

To establish which combination of these measures best explains participants' judgments we entered them all into a competitive, stepwise, model selection procedure. We used all the data for the model selection. As before we fit mixed effect models with random effects for devices within participants. The independent variables were first z-scored meaning that the final beta weights can be interpreted as percentage increase in assignments to the chain for a 1 standard deviation increase in the value of each independent variable. We entered the following predictors:

Intercept: Positive value captures overall preference for chain, negative for fork.

cor(t_{SA}, t_{SB}): The correlation between the delays $t_{SA}^{1:k}$ and $t_{SB}^{1:k}$.

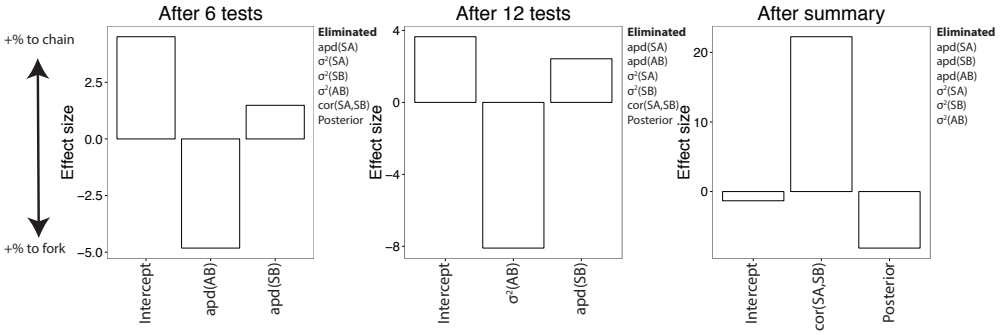


FIGURE 6.16: The models resulting from stepwise model selection to all 1040 1st, 2nd and final judgments in Experiment 8. Plots show the selected predictors' fixed effects (i.e. the β values) in descending order of t value. All predictors were z-scored, and the dependent variable was centred (so that a prediction of 0 corresponded to assigning 50% to the chain and 50% to the fork). Thus, effect sizes are interpretable as differences in percentage assigned to the chain moving one standard deviation up on the independent variable.

$\boldsymbol{\sigma^2(t_{xy})}$: The variance of the inter-event timing between activation of components x and y in the tests performed so far. We entered the variance for each inter-event interval (i.e. $\sigma^2(t_{SA})$, $\sigma^2(t_{SB})$ and $\sigma^2(t_{AB})$).

$\mathbf{APD}(\boldsymbol{t_{xy}})$: Average pairwise difference. A sequentially computed proxy for variance. The difference in activation time on current test compared to previous test for example t^k_{xy} and t^{k-1}_{xy} summed up over tests $1:K$. E.g. for t_{AB} after six trials this is $\text{APD}(t_{AB}) = \sum_{k=2:6} t^k_{AB} - t^{k-1}_{AB}$. As with the variance, we entered the APD for each inter-event interval.

Posterior: The posterior probability of a chain according to Delay_I

Figure 6.16 depicts the models selected by the stepwise procedure for the first, second, and third judgments respectively. In all three cases, 2 of the 8 predictors were chosen and the rest eliminated. The chosen predictors were similar for the first two judgments but quite different for the final judgment.

For the first judgment — after 6 tests — participants assigned fewer points to the chain (and more to the fork) if there was high apparent fluctuation in t_{AB}, measured by comparing each test to the previous (i.e. $\text{APD}(t_{AB})$). Fluctuation in t_{SB} was also selected but had a smaller effect in favour of the chain. The fact that the intercept is $\gg 0$ is also suggestive of a baseline preference for the chain that could be overturned by high $\text{APD}(t_{AB})$ or low $\text{APD}(t_{SB})$. The Bayesian posterior was not selected as part of the final model.

For the second judgment — after seeing all the evidence — we see a similar pattern but this time the actual variance of t_{AB} is selected rather than its sequential proxy. Again there is a baseline preference for the chain and a weaker influence of APD(t_{SB}).

For the final judgments — made after seeing the visual summary — we find a different pattern. Now the correlation between t_{SA} and t_{SB} dominates the selected model, so much so that there there is a significant negative relationship with the Delay_I posterior.

We report the correlation between all the predictors in Appendix B.

Summary

In sum, these additional analyses suggest that participants had an initial preference for the chain which was modulated based on their perception of variability in t_{AB} and, to a lesser extent, in t_{SB}. This is consistent with the idea that many began with an (order-driven) preference for the chain which they could gradually reject if their experienced delays were highly variable under the chain hypothesis. After the visual summary was available, judgments shifted to reflect predominantly the more reliable, but harder to compute, predictive relationship between t_A and t_B — $\text{cor}(t_{SA}, t_{SB})$ — which "popped out" visually when viewing the summary timeline (Figure 6.14b).

6.7 General Discussion

In our first two experiments, we found that people were adept at using order information to make judgments about causal structure, based on a single trial (Experiment 6) and by integrating the information from several observations (Experiment 7). We found that participants generally made the non-simultaneity assumption embodied by our Order_N model, but also distributed likelihood in a way consistent with a preference for similar causal delays within each device. Additional variability could be explained as resulting from uncertainty about how causes combined in the Collider (common effect) structure, some additional sensitivity to the precise event timings, and some degree of conservatism in evidence integration. In Experiment 8, we removed the order cues. In this setting, participants were able to use the variability in the event timings alone to distinguish between a chain and a fork structure. To our knowledge, this is the first time this has been shown experimentally. We now discuss these results more broadly and propose some future directions.

6.7.1 Nonsimultaneity and simultaneity

Like Burns and McCormack (2009) and McCormack et al. (2016), we found the large majority of our participants made judgments in line with a non-simultaneity assumption. This means they considered events that occurred at the same time to be inconsistent with one being caused by the other. However, in parallel they had a preference *for* simultaneity among events that shared a common cause (forks), or effect (colliders), judging it at least as likely that these "common" events will occur simultaneously as one occurring earlier or later. From a continuous time perspective, the probability of perfect simultaneity given variability is strictly zero, while nonsimultaneity, if untied to any particular delays, covers the rest of the space. However, human perception does not have infinite temporal precision. Assuming some perceptual uncertainty, apparent approximate simultaneity is the most likely outcome, with likelihood falling away the greater the perceived discrepancy between the outcomes in either direction.

6.7.2 Causal time perception

We looked at only a narrow range of time intervals in the current studies, with trials never lasting more than around three seconds. Weber's law (1834) states that perceptual estimation errors normally grow in proportion to the quantities involved. However, this is known to break down for short (< 1 second) intervals which are tracked differently by the brain (Karmarkar & Buonomano, 2007). For short intervals, existing causal beliefs have been shown to shape, or distort, perception (Buehner & Humphreys, 2009; Haggard et al., 2002), sometimes even leading to reordering of a surprising series of events to a more "normal" causal order (Bechlivanidis & Lagnado, 2016). This suggests that at this temporal grain, experience is still somewhat under construction (Dennett, 1988), scaffolded by preexisting expectations about causal structure. This also suggests an explanation for why participants in the current experiments sometimes seemed to retain some preference for devices that should have been ruled out (as captured by our ϵ parameter in Experiment 7).

Having formed an impression that a device has a certain structure, someone might easily misperceive a subsequent observation as consistent even if they would usually consider it inconsistent with that structure. This might occur more often if the distortion required to make it consistent is very small. In particular, the simultaneous events that people considered to rule out causation most of the time, might also have been susceptible to being perceived as occurring in the expected causal order. These sorts of effects

are not captured by our Order and Delay models which work at Marr's (1982) computational level, and are intentionally scale invariant. However, an interesting project would be to construct a cognitive model that exhibits these patterns. Related to this, a fundamental reason to expect different learning at different timescales comes from the so-called "now or never bottleneck" (Christiansen & Chater, 2016) inherent to experiencing events in real time. When observing closely spaced events, there is little time for explicit comparison of possible structures, or to do anything much beyond constructing an impression of what happened or measuring how wrong your prediction was. Reasoning about relationships between events that are separated by minutes or hours is likely a very different process, as there would be far more time to explicitly reason about and compare hypotheses.

6.7.3 Modality

In the current tasks we looked only at the visual modality. However, it could be that other modalities are even better at inferring patterns in time, audition being an obvious example. Humans (and many other animals), have a finely developed ear for patterns in time and pitch; allowing us to hear and quickly internalise even complex rhythms and melodies (London, 2012). Furthermore, the brain can detect an auditory pattern amongst noisy background and even decompose it into its constituents elements (i.e. distinguishing the different instruments in a band). It seems plausible that we evolved these capacities in part to support the search for the reliable patterns in nature that are often clues to its underlying structure (Sloman, 2005). Supporting the notion that the visual modality is better at spotting spatial rather than temporal patterns, we saw that participants were able to make much stronger judgments in Experiment 8 once they saw a visual summary. The summary replaces temporal distance with spatial distance, and suddenly the reliability of t_{AB} pops out clearly (as in Figure 6.14b). During the trials themselves, this realisation depended on effortful memorisation and comparison across observations.

6.7.4 Conjunctive influence

In Experiments 6 and 7, we instructed participants that the Collider structure was conjunctive — that is, it required both of its causes to activate before the effect would activate. We also included a comprehension question to check that participants had understood this. Nevertheless, around a quarter of participants across both experiments,

appeared to treat the Collider as disjunctive (or at minimum not rule out that it was capable of behaving disjunctively sometimes), assigning nonzero probabilities to the Collider even after observing clips where one of its cause components occurred after the effect, or nonzero likelihoods for the Collider to patterns with only one cause occurring before the effect. This suggests that people default to the disjunctive assumption so strongly that it can either overrule instructions, or fill in if the instructions were forgotten (cf. Lu et al., 2008; Lucas & Griffiths, 2010; Yeung & Griffiths, 2015).

Additionally, people might have struggled to make sense of the idea of a conjunction in the context of the abstract tasks they were solving. Indeed, formalising the conjunction for our Delay models forced us to think about what would be a plausible mechanism. Concretely, we assumed that the earlier-arriving causal influence waited around in a buffer for the latter to arrive. However, it would have also been plausible to assume that the two causes have influences that must (at least approximately) coincide in their arrival time in order for a threshold to be reached that triggers the effect. Additionally, people might find conjunctive influence more natural in situations where at least one of the causal relationships has a sustained or continuous effect (e.g. so that the second event simply tips the level of influence over a threshold that causes the activation of the effect). In general, participants were more uncertain about devices where the impact of the evidence depended on assumptions about how the Collider worked. Given the ambiguity about the exact way in which the Collider worked, participants' increased uncertainty for situations involving these cases may be considered a rational response.

6.7.5 The blessing of variability

Our experimental design highlights an interesting and counterintuitive property of temporal causal inference. Unreliable systems can actually be simpler to uncover. The more unreliable the timings of the events are, the more frequently revealing order reversals will occur, and the more a learner can rely on simple qualitative Order inference. A similar principle applies in the absence of revealing order information. It is actually the variability in delays that provides the signal that our Delay models use to infer the generative causal structure. If the causal delays are perfectly reliable it becomes impossible to distinguish between the order-consistent structures based on their timing.[17] This has interesting parallels to the case of learning from contingency information. In a

[17] Assuming you do not have a prior expectation about the lengths of the different delays. Of course, structures could still be distinguished without variability if you know how long the links should take to work.

deterministic system, chains and forks are indistinguishable from contingencies because both effects always covary with their root cause. However, they can be covariationally distinguished in various settings provided the relationships are at least a little unreliable (Bramley, Dayan, & Lagnado, 2015; Fernbach & Sloman, 2009).

6.7.6 Toward a process model

Some participants in Experiments 6 and 7 formed preferences for structures that rendered causal delays more similar on average across connections and clips (reflected by the shift across clips 2 to 6 in Experiment 6, and the few individuals better described by our time-sensitive Delay_P than our qualitative Order models). Additionally, participants' distribution over qualitative patterns was highly consistent with a preference for equal delays. Judgments in Experiment 8 were consistent with the proposal that people tend to "see" the evidence through the lens of one causal model at a time (Bechlivanidis, 2015), becoming more likely to switch if observed events are sufficiently hard to accommodate under this presumptive structure (Bonawitz et al., 2014; Bramley, Dayan, et al., 2017). Since seeing several events that always occur in the same order *ceteris paribus* is most naturally perceived as a chain, participants may have begun the problems in Experiment 8 with a sense of watching a causal chain, which could be gradually overturned in the cases where there was another more predictive perception available (of the device as a fork). More generally, by pulling these ideas together, we get a picture of temporal causal structure learning in which learners have an initial impression of causal structure based on event order (Lagnado & Sloman, 2006) but are capable of refining this as they observe more evidence about the system and consider what structural changes from this default might make the event times more predictable.

6.7.7 Building richer causal representations

While CBNs provide our current best framework for building theories about causal cognition, they are not rich enough to explain central aspects of causal cognition such as mechanism knowledge and mental simulation (Mayrhofer & Waldmann, 2015; Sloman & Lagnado, 2015; Waldmann & Mayrhofer, 2016) or to ground everyday causal judgments (Gerstenberg et al., 2015). People's causal representations almost certainly lie somewhere in between a compact statistical map (a CBN) and a scale model of the physical world. We can often get away with treating detailed mechanisms as black boxes (Keil, 2006), but we still need our representation to help us choose when and where to act

in the world. Thus, it seems necessary that people's representations sometimes include expectations about delays between causes and effects. Of course our causal representation of the world is rich in space as well as time, with detailed knowledge of mechanisms likely to be intertwined with delay expectations. Our generative Delay models represent a step toward capturing the ways in which human causal cognition goes beyond statistical contingencies.

6.8 Conclusions

In conclusion, this chapter showed in three experiments that people form clear and sensible beliefs about causal structure based on temporal information. We can capture people's inferences with a combination of qualitative order-based, and generative delay-based inference models. Participants were able to use the order in which events occurred to narrow in on candidate causal structures, and within these, favoured those that rendered the causal delays more similar and more predicable. Going beyond order patterns, we showed that people can also use interval variability alone to identify whether a structure is a chain or a fork, and proposed how participants might achieve this while "seeing" the evidence through the lens of one hypothesis at a time. These results contribute to understanding of the role of time in causal learning and representation, showing that just as time is inherent to our experience of the world, it is integral to our causal models of the world.

Chapter 7

Intervening in time

"*One should know when to act and when to refrain from action.*"

— DALAI LAMA

In a temporally continuous world, using interventions to uncover causal relationships requires good timing. For instance, it is hard to tell whether a new medication is effective if you take it on top of others, or just as you start to feel better. Likewise, it is hard to tell whether a new law lowers crime if it is introduced just after other reforms or before a major election. Such inferences, having to do with delayed effects of actions and a changing causal background, can be particularly tough in dynamic systems where feedback loops make prediction difficult (Brehmer, 1992). In short, for interventions to be effective tools for unearthing causal structure it is important to time and locate them carefully, while paying close attention to the time course of surrounding events.

Learning by associating actions with surrounding events and stimuli was traditionally studied through free operant conditioning (e.g. Mackintosh, 1983; Skinner, 1963). In free operant conditioning, subjects can perform an action at will, and subsequently receive rewards or punishments on a specific delay schedule determined partly or entirely by the performance of these actions. The result is positive or negative reinforcement of actions as well as their association with other paired stimuli (Estes, 1948). In recent years, new interpretations of learning in these tasks have have been proposed, based on the idea that subjects gradually learn a causal model of the task, based on regularities in the reinforcement delays and latent trial structure (Gershman, Jones, Norman, Monfils, & Niv, 2013; Gershman & Niv, 2012; Greville & Buehner, 2007, 2010). This is consistent with a large amount of recent work suggesting people are adept at inferring causal structure from interventions (Bramley, Dayan, et al., 2017; Bramley, Dayan, & Lagnado, 2015; Bramley,

Lagnado, & Speekenbrink, 2015; Coenen et al., 2015; Sobel & Kushnir, 2006; Steyvers et al., 2003) and based on temporal information (Bramley, Gerstenberg, & Lagnado, 2014; Bramley, Gerstenberg, Mayrhofer, & Lagnado, submitted; Buehner & May, 2002, 2003, 2004; Buehner & McGregor, 2006; Greville & Buehner, 2007, 2010; Lagnado & Sloman, 2006; Lagnado & Speekenbrink, 2010). Nevertheless, the shift toward thinking about free operant learning in terms of causal model induction opens up a large new theory space. Learning a single action–outcome pairing is only the tip of the iceberg. The more general problem lurking beneath is that of learning latent causal structure from actions and events in continuous time. What should we learn when our actions are succeeded by multiple events, some expected others unexpected? And *when* should we act, if we want to "condition ourselves" pro-actively, to form associations representing the true causal relationships but not spurious or coincidental ones? Some work has looked at causal learning from point events in time (Deverett & Kemp, 2012; Lagnado & Speekenbrink, 2010; Pacer & Griffiths, 2015; Rothe, Deverett, Mayrhofer, & Kemp, 2016) but none, to our knowledge, on the role of interventions in this context. This chapter describes an experiment that takes a small step toward exploring this general problem.

The structure of the chapter is as follows: We first describes the learning problem we focus on, then review recurring ideas from the previous chapters using these to motivate an experiment and some modelling.

7.1 The learning problem

We explore the general problem of how people infer causal structure from interventions and subsequent patterns of events (component activations). We focus on identification of the causal structure of mystery "devices" made up of several components that can exhibit multiple instantaneous events, or *activations*, over a continuous period, as in a point process (Kingman, 1993; Norberg, 1986).[1] Where components of these devices are causally related, each activation of the parent component will normally cause a single subsequent activation of an effect component after a delay. We assume that these delays are Gamma distributed with an average delay length μ and some variability governed by a shape parameter α (see Equation 6.3). We restrict our focus to situations with no spontaneous component activations, but where causal relationships only work stochastically (e.g. with probability w_S). Any pair of components can be connected in

[1] A point process is a type of random process for which any single realisation consists of a set of isolated points, such as events of negligible duration located on a timeline or trees located in a forest.

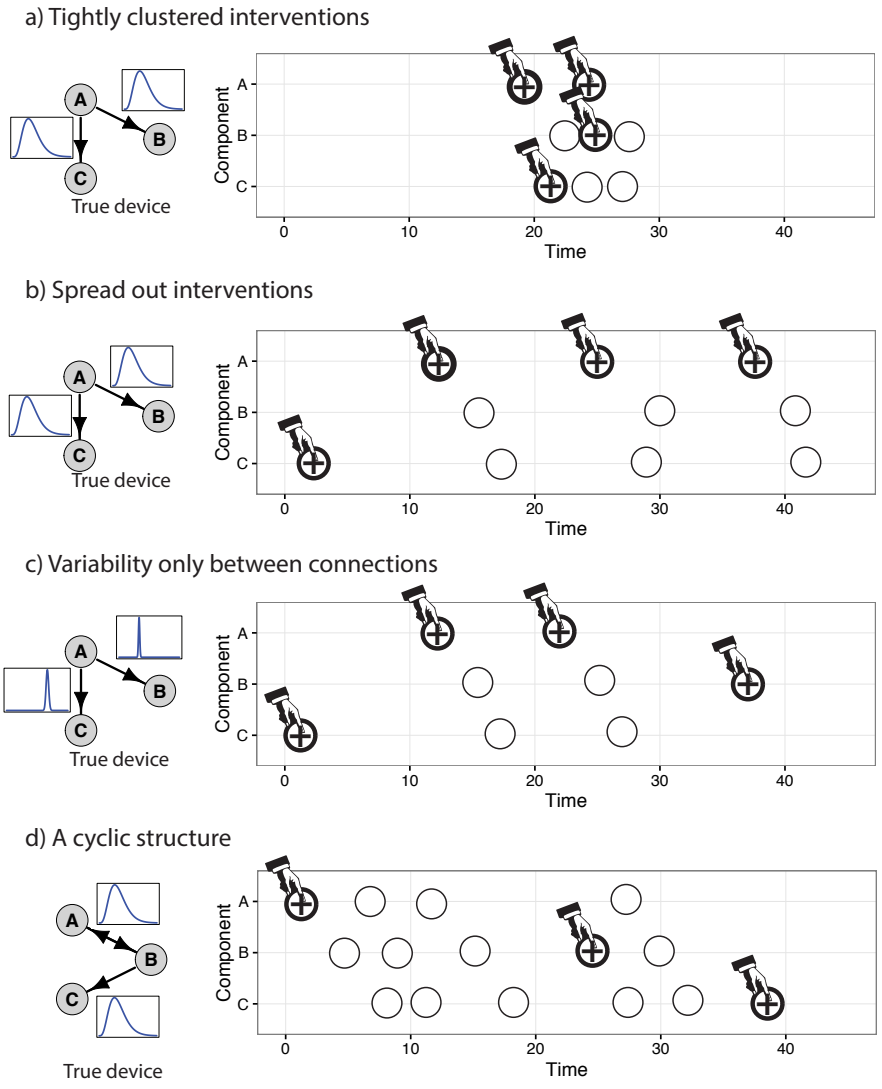


FIGURE 7.1: Examples of using real-time interventions to infer causal structure. a)–d) show timelines, each with a row for each component of the causal system. Circles with a "+" symbol and incoming hand icon are interventions. Plain circles are activations. The true generative causal model is depicted on the left hand side, with subplots for each edge depicting the delay distribution (each with a mean of 1.5 seconds).

either, neither or both directions, but components cannot be connected to themselves. This results in a hypothesis space $m \in \mathcal{M}$ of 64 possible structures for devices made up of three components, or 4096 for four components. Furthermore, we assume learners can intervene on the devices by causing any component to activate at any moment of their choosing. Such interventions have the same causal effects as non-interventional events but need no causal explanation themselves, analogous to the interventions on CBNs as modelled by Pearl's Do[.] operator (e.g. see Section 3.1). Figure 7.1 shows a number of examples of such interactions, taking place over a short period, with further explanation below.

7.1.1 Interventions as structuring learning in time

In Chapter 5, we saw that people choose interventions to support their limited and incremental learning trajectories. In the CBN context, this meant focusing on one part of a problem at a time rather than trying to learn about everything at once. However, the continuous-time case has different sources of complexity. Seeing the effects of one's interventions play out in time provides rich information, making causal model inference easier in some respects. However, the other side of the coin is that there are no completely independent trials in continuous time. In general, in observing a continua of events, one cannot rule out that something that happened earlier is still exerting its influence, or that an effect is yet to reveal itself by the time one stops watching. Fortunately, interventions provide anchor points, that we can be sure are not effects of anything else, affecting the future but not the past (Lagnado & Sloman, 2004). Thus, one way that interventions can revive the notion of discrete trials is by creating a trial structure in time. By waiting long enough between interventions to be confident effects have dissipated, we can turn an undifferentiated event stream into something more structured and informative about causality. Figure 7.1a gives an example of interventions on a Fork device that, intuitively, are not well chosen. The learner performs four interventions in close succession and experiences four outcomes. However, it is hard to attribute causal responsibility for these activations to particular prior activation or intervention events, since there are so many similarly plausible candidates. In contrast, Figure 7.1b shows an example where the interventions are spaced more widely, intuitively resulting in less ambiguous information.[2] To date, no one has looked at how people select when to intervene when learning a causal system in continuous time.

7.1.2 Variability and positive testing

In Chapter 6, we found that people were adept at learning causal relationships from event order, and other things being equal, they favoured causal structures that rendered the causal delays more similar. When a cause–effect delay was internally variable — i.e. the same relationship would sometimes take longer and sometimes shorter — learners could also use inter-event variability and correlation to distinguish direct and indirect causation, although this required considerably more evidence than the inference based simply on order or expectations about delay lengths. We noted that two of these uses of

[2] In experimental psychology we often achieve such approximate independence by including distractor tasks between trials, counterbalancing the order in which we present stimuli, and by repeating tests "between-subjects".

temporal information — order reversals and correlation between inter-event intervals — depend on there being some variability in delays. The same is true in the current context. In Figure 7.1b, the Fork structure is revealed partly by internal variability. Interventions on A lead to cases where B precedes as well as succeeds C allowing (attentive) learners to infer the device is neither an $A \rightarrow B \rightarrow C$ Chain nor an $A \rightarrow C \rightarrow B$ Chain. In contrast Figure 7.1c shows an example where delays are internally reliable but vary between the connections. This is shown by the strongly peaked delay distributions with differing means — shown in blue, next to edges in the model on the left with $t_{A\rightarrow C}$ reliably longer than $t_{A\rightarrow B}$. Here, repeatedly intervening on the A component does not help the learner distinguish the Fork from the Chain hypotheses and the learner must intervene on both B and C to get clear evidence that the device is a Fork. Whether people make use of such variability information in a free interventional context is as-yet unexplored.

In discrete-trial contexts, we saw people had a preference for interventions on root components. While often not the most globally informative choice, this was an effective way of assessing the adequacy of their current working hypothesis. It could be that we see a similar pattern of positive testing in the continuous-time case. This could be especially useful when there is variability within the cause–effect delays since the causal order of subsequent events can often be inferred based on attention to order reversals and variability and correlation among delays. Indeed, the evidence in Experiments 6, 7 and 8 can all be seen as that resulting from a learner repeatedly activating the root component of a causal device and relying on time to reveal the patterns of forward connectivity.

7.1.3 Causal cycles

The preceding chapters focused on acyclic causal systems where causal influences can flow in only one direction, never revisiting the same component. This was partly in virtue of the conceptual and mathematical convenience afforded by the acyclic CBN framework. However, our understanding of many physical, biological, social and economic aspects of the world are inherently cyclic (Malthus, 1888). Furthermore, people frequently report causal beliefs that include cyclic relationships when allowed to do so (e.g. Kim & Ahn, 2002; Nikolic & Lagnado, 2015; Sloman et al., 1998). While there are ways of adapting the CBN formalism to capture cycles (Rehder, 2016), these either simplify the problem to influences between fixed time steps (e.g. Dean & Kanazawa, 1989; Rottman & Keil, 2012), or simplify the representation by modelling dynamic subnetworks only by their

equilibrium distributions (e.g. Lauritzen & Richardson, 2002; Rehder, 2016). By focusing on continuous event streams and developing the continuous-time causal representation introduced in Chapter 6, it is possible to explore how people learn about continuous-time dynamic cyclic systems.

Not much is known about how people learn in cyclic causal structures (although see Deverett & Kemp, 2012; Rothe, Deverett, et al., 2016), with most work in dynamic causal contexts focusing on control rather than structure learning (e.g. Brehmer, 1992; Osman, 2011). From a formal perspective such systems are often fundamentally hard to predict even if you know how they work (e.g. Kushner, 1967). Feedback loops often lead to sensitive dependence on initial conditions, with large changes in behaviour stemming from small perturbations (Gleick, 1997), explaining why we cannot predict the long term behaviour of dynamic systems like the weather or the economy much above chance. Thus, a reasonable hypothesis is that complex dynamics make cyclic causal structures harder to learn than acyclic ones, potentially requiring different interventions or inferential tools for success. Figure 7.1d gives an example of interventions on a cyclic causal system (assuming that the connections work 90% of the time). Interventions initialise looping behaviour because of the bidirectional relationship $A \leftrightarrow B$ (e.g. $A \rightarrow B \rightarrow A \rightarrow B \ldots$) leading to many subsequent activations of both the loop components and the output component C, continuing until either the $A \rightarrow B$ or $B \rightarrow A$ connection fails. Based on simply looking at the timeline it seems likely that it will be easier to identify which components are either directly involved in cycles, or outputs from cyclic components (due to their recurrent activations), but harder to identify the exact causal relationships (e.g. whether it is A or C that causes B in this example, since both tend to recur shortly before B).

In order to look formally at learning in cyclic and noncylic systems we must extend our normative framework to handle cases where components can exhibit multiple activations.

7.2 Modelling continuous time causal learning

7.2.1 Normative inference

In Chapter 6 we treated each discrete trial as a data point in which all components of a device would activate (e.g. $d_i = \{t^i_A, t^i_B, t^i_E\}$). Since we are now interested in situations that lack this discrete trial structure, and components can have more than one activation

within a trial, we require different notation. We now consider trial data $\mathbf{d}$ to be made up of all activations of all components within an observational window $[\tau_{\text{start}}, \tau_{\text{finish}}]$:

$$\mathbf{d}_{[\tau_{\text{start}},\tau_{\text{finish}}]} = \left\{ d_X^{(1)}, \ldots, d_Y^{(n)} \right\} \tag{7.1}$$

in which events are indexed in chronological order with subscripts denoting their "location" (e.g. which component activated). Data $\mathbf{d}$ is conditional on any interventions $\mathbf{c}$ performed prior to τ_{finish}, e.g.:

$$\mathbf{c}_{[-\infty,\tau_{\text{finish}}]} = \left\{ c_X^{(1)}, \ldots, c_Y^{(n)} \right\}. \tag{7.2}$$

For instance, one might interact with a causal device for 5000 ms, performing interventions on components A and B: $\mathbf{c}_{[-\infty,5000]} = \{c_A^{(1)} = 100, c_B^{(2)} = 1200\}$, and observing two activations of C: $\mathbf{d}_{[0,5000]} = \{d_C^{(1)} = 1500, d_C^{(2)} = 2800\}$.[3]

Normative structure inference thus involves computing the likelihoods $p(\mathbf{d}|m;\mathbf{c})$ and using these to update a prior $P(M)$ to a posterior $P(M|\mathbf{d};\mathbf{c})$. An immediate issue with performing Bayesian structure inference in this setting is that, even for a single candidate model, there are likely to be multiple potential paths of *actual causation* that could have given rise to observed data (Halpern, 2016; Halpern & Hitchcock, 2011; Halpern & Pearl, 2005). For example, if the true structure is a $A \rightarrow C \leftarrow B$ Collider, the data above might be produced in two ways. A could have caused the first activation of C and B the later ($c_A^{(1)} \rightarrow d_C^{(1)}, c_B^{(1)} \rightarrow d_C^{(2)}$). Alternatively, A could have caused the later activation of C and B the earlier ($c_A^{(1)} \rightarrow d_C^{(2)}, c_B^{(1)} \rightarrow d_C^{(1)}$).

For sufficiently small numbers of events, it is possible to enumerate all such possible paths of actual causation under each model $m \in \mathcal{M}$ (we call these $\mathbf{z} \in \mathbf{Z}_{\mathbf{d}}^m$) summing over them to get a marginal likelihood for $p(\mathbf{d}|m,\mathbf{c})$. For a large number of events this becomes intractable[4] but we were able to compute the posteriors in this manner for the data from the current experiment, resorting only in rare cases to an approximation. Details on how we compute the likelihoods and the posterior probabilities $P(M|\mathbf{d},\mathbf{w};\mathbf{c})$ are provided in Appendix B.

[3] We assume that there were no interventions before τ_{start}.

[4] Inference can proceed in these settings by modelling causes as exerting temporary changes on the *rate* of their effects' occurrence over time (see Pacer & Griffiths, 2011, 2015).

7.2.2 Incremental construction heuristics

A core theme of this thesis is the idea that people construct their causal models incrementally, accommodating new evidence by making local changes to a single global model (e.g. see Chapter 5, and Bramley, Dayan, et al., 2017; Bramley, Dayan, & Lagnado, 2015). This idea seems particularly applicable to the continuous-time context, where normative inference is tough and the evidence, by its nature, arrives continuously.

Chapter 5 proposed that people update their causal models in order to accommodate recent evidence. We compared *simple endorsement* — a heuristic that adds links to explain effects directly, with no consideration of the existing model — with *Neurath's ship* — a more sophisticated scheme in which connections were reconsidered in the light of the rest of the existing model. In the continuous-time context, the problem of accommodating the latest evidence comes up every time an event occurs. A simple heuristic strategy would thus be to attribute a cause to each new event, irrespective of the pre-existing structural beliefs, while a Neurath's ship style strategy would be to add new links only if the existing structure belief cannot explain the new evidence. In terms of diagnosing the likely causes of each new activation, we can look to Chapter 6. It explored inference from event order alone, as well as sensitivity to inter-event intervals, finding that people made use of order for ruling out models, but also favoured models that rendered observed event timings more consistent overall.

Combining levels of existing-model sensitivity and timing sensitivity suggests several potential heuristics for the current setting. These are based on the idea of adding to or adapting single model b as events are experienced. The result in each case is a single structural belief that can evolve over time (we write $\mathbf{b} = \{b^{(0)}, \ldots, b^{(n)}\}$, where the sequence of belief indices correspond to the event indices in $\mathbf{d}$). We propose several heuristics differing in the sophistication with which they diagnose the cause of each newly experienced effect $d^{(i)}$, and so adapt $b^{(i-1)}$ to form $b^{(i)}$:

1. **Add most recent** Each time an event is observed, this heuristic simply attributes it to the most recently preceding event at any other component (either the most recent intervention in $\mathbf{c}$ or activation in $\mathbf{d}$). If $b^{(i-1)}$ does not contain an edge from the location of this preceding activation to the location of the current activation, it adds this to $b^{(i-1)}$ to make $b^{(i)}$. Figure 7.2a gives an example of this. Starting from $b^{(i-1)}$ with a single $D \rightarrow B$ connection, the heuristic connects A to B upon observing B's activation, and then B to C when C activates shortly after.

2. **Add most likely** This heuristic is like *most recent* except that, instead of attributing activations to the most recent preceding event, it attributes to the previous event that was most likely to bring about the observed activation, given expectations about the true causal delays (e.g. knowledge about mean the mean μ and variability α of true delays). Again, if there is not already an edge from the location of this most likely cause running to the location of the current activation, this heuristic adds this to $b^{(i-1)}$ to make $b^{(i)}$. Thus this model captures the kind of delay sensitivity in causal attribution that has been thoroughly demonstrated in the literature on delay based attribution (Buehner & May, 2002, 2003, 2004; Buehner & McGregor, 2006; Greville & Buehner, 2010). However these attributions are still done irrespective of existing connections in $b^{(i-1)}$. Figure 7.2b gives an example of this. Here, C's activation time is most consistent with its being caused by the intervention on A, thus the model adds an $A \to C$ connection, rather than $B \to C$ connection, going into $b^{(i+1)}$.

3. **Add more likely** This heuristic is like *most likely*, except that rather than automatically adding a connection if the most likely explanation is not already connected in $b^{(i-1)}$, it first checks if there is already an adequate explanation in the current model $b^{(i-1)}$. Concretely, it compares the likelihood of the most likely explanation that is already a cause in $b^{(i-1)}$, to the most likely explanation *overall*. Where these differ, it only adds a new connection if the best overall explanation is substantially *more likely* than the best existing explanation in $b^{(i-1)}$ (where this is determined by passing some predefined significance level). Thus, this model embodies the conservatism discussed in Chapter 3 and *Neurath's ship* style model-based inference (Chapter 5), where the learner first tries to explain the effect with their existing structure, and only adds a new link if it provides a much better explanation of the data. Figure 7.2c gives an example of this idea. Unlike *most likely*, this heuristic does not add an $A \to B$ connection going into $b^{(t+1)}$ because B's activation can be explained well enough by the existing connection $D \to B$. While $p(t_{c_A^{(j+1)} \to d_B^{(i)}})$ is slightly more probable than $p(t_{c_D^{(j)} \to d_B^{(i)}})$ the difference is not substantial enough to warrant the addition of another connection.

4. **More likely + pruning** This heuristic is the same as *more likely*, except that it performs an additional pruning step, allowing it to remove connections if they appear to be repeatedly failing to work. Not only does this heuristic take the current model into account when interpreting evidence, it also keeps a count for each connection in b, of previous activations (and interventions) of the cause component

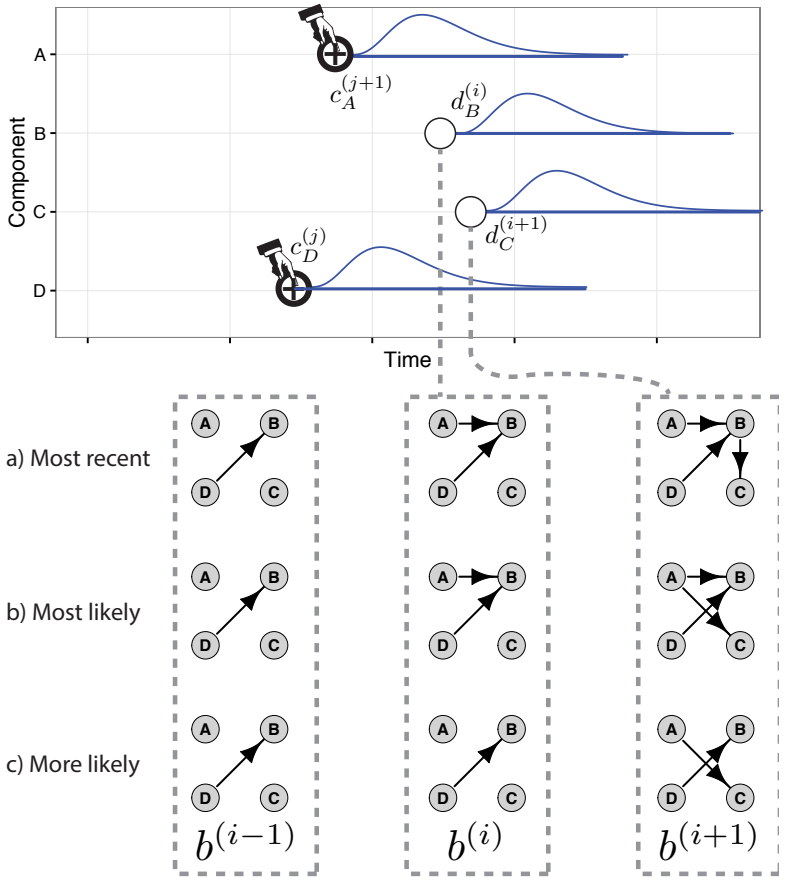


FIGURE 7.2: An example in which the proposed heuristics make diverging model construction decisions. $b^{(i-1)}$ shows the learners belief from prior to the period depicted in the timeline plot. After observing an event at component d_B^i, the models update $b^{(i-1)}$ to form $b^{(i)}$ and then after observing d_C^{i+1}, update again to form $b^{(i+1)}$. Blue lines indicate the probability density for the likelihood of cause–effect delays starting from each event, used to determine the *most likely* cause of each event, and whether it is sufficiently *more likely* than any existing causes.

and previous activations of the effect, calculating the binomial probability (given w_S) of observing k successful activations of each effect in the model given that there have been n activations of the cause, every time a new event is observed.[5] Thus, unlike 1–3, this heuristic is able to remove connections as well as add them.

The key questions of interest in the current experiment are:

1. How people interact with the devices during learning. How do they distribute their interventions across trials and across the components of the system?

2. How do these choices affect the information they receive about the true connections and their consequent judgment accuracy?

[5]There are a range of complications here, e.g. some events have multiple causes or might have effects that have not happened yet. Finding a clean way of dealing with these issues is a work in progress.

3. In what way are interventions and judgments affected by true structure of the device and the nature of the true cause–effect delays?

4. Will we see patterns of incremental model construction similar to Neurath's ship in the discrete time context?

7.3 Experiment 9: Intervening in time

We tasked participants with performing interventions over a 45 second interval with the goal of identifying the causal relationships between the components. We focused on simple generative interventions (e.g. clicking on components to activate them) and a mixture of cyclic and acyclic devices. We were interested in whether learners would spread their interventions out or bunch them together, as well as how they would distribute them over the components. Given the role of variability in distinguishing direct and indirect causation in Chapter 6, we were also interested in how learning and intervention choice was affected by the nature of the inter-event delays in the causal devices, whether they were reliable or variable, and whether this variability was *within* connections (e.g. Figure 7.1b) or *between* them (e.g. Figure 7.1c). In general, we expected that reliably similar delays between edges would make the devices easier to learn. For devices exhibiting variability *within* connections, we hypothesised that learners would use repeated positive testing to distinguish direct and indirect relationships. For cases with variability only *between* connections, we expected lower performance, and reliance testing each component separately, since both order and variability are unreliable guides to cause. We were also interested in how well learners could learn different types of devices, having hypothesised that they would find cyclic devices harder than acyclic ones. Finally, we were interested whether learners' behaviour would be consistent with the principles identified in earlier chapters, namely confirmatory testing, and incremental construction of a single global model.

7.3.1 Methods

Participants

Sixty participants (24 female, aged $M \pm SD$ 32.9 ± 10.0) were recruited from Amazon Mechanical Turk so that 20 performed in each of three conditions. The task took

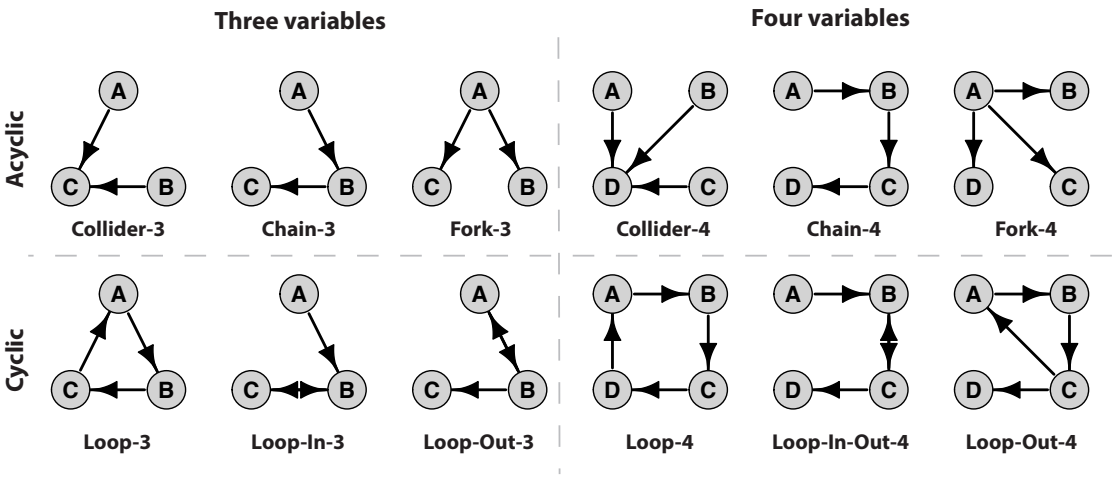


FIGURE 7.3: Generative models tested in Experiment 9

around 20 minutes, and participants were paid between \$0.50 and \$3.20 depending on performance ($M \pm SD$ \$2.06 $\pm$ 0.39).

Stimuli and design

Each participant interacted with 12 causal devices over 45 second trials. We included a range of acyclic and cyclic devices including a 3- and 4-variable Collider, Chain, and Fork as in previous chapters. We label these with suffixes denoting the number of variables involved (e.g. *Collider-3* for the three variable Collider see Figure 7.3). We also included six novel cyclic devices. We label these depending on whether they have a connection feeding *In* and/or *Out* of the cyclic subnetwork (e.g. *Loop-3-In* has a connection going from A, *in*to the cyclic $B - C$ subnetwork, see Figure 7.3). We will refer to the components involved in the cyclic relationships as *loop components*, any components that feed activations into a cyclic group (e.g. A in Loop-In-3) as *input components* and any whose activations are emitted out from a cyclic group (C in Loop-Out-3) as *output components*.

In all devices, causal relationships worked 90% of the time ($w_S = 0.9$) and there were no background activations ($w_B = 0$). The average delay between the activation of a cause and an effect was 1.5 seconds. However, we examined three delay conditions between subjects that differed in the extent and nature of the delay variability:

1. **Reliable** In this condition, the activation of a cause component led to the activation of its effect with little variability. Concretely, all delays were gamma distributed with mean $\mu = 1500$ms and shape $\alpha = 200$ (Figure 7.4, full line).

2. **Variable-within** In this condition, delays had the same mean but were much less reliable with shape $\alpha = 5$ (see Figure 7.4, dashed line).
3. **Variable-between** In this condition, each causal relationship was assigned a delay at the start of the trial, drawn from the same distribution as the *variable-within* condition, but would then stay the same throughout the problem. Thus, in this condition, delays varied considerably between relationships but were reliable within a single relationship, taking the same amount of time to work every time they were tested.[6]

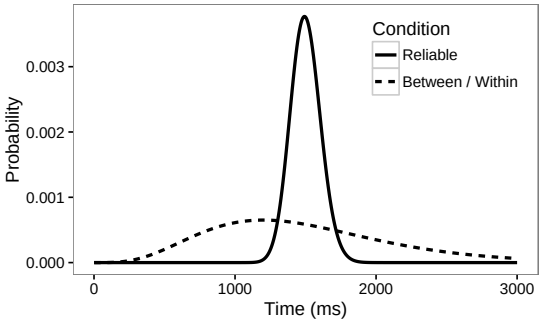


FIGURE 7.4: Delay distributions by condition.

The differences between conditions affect what can be inferred about the structures. In the *reliable* condition, the absolute interval between two events is a strong indication of whether the former is a plausible cause of the latter. For example, suppose you intervene on A at $t = 0$, then observe B at $t = 1500$ and C at $t = 3000$ and you know that either A or B caused C. In the *reliable* condition, the timings are strong evidence that B not A caused C (the likelihood ratio is $\approx 10^{26}$). However, in the *variable-within* condition, it is only weak evidence (the likelihood ratio is ≈ 9). Repeatedly testing the same component is more useful in the *variable-within* than *variable-between* condition, primarily because it can lead to revealing order reversals.

Materials and procedure

The task was programmed in javascript, hosted online, and can be tried out here `https://www.ucl.ac.uk/lagnado-lab/el/nbt`.

Each device was represented by several grey circles on a white background with light grey boxes marking the potential locations of edges (see Figure 7.5a). Participants were told that the circles were components of a causal device. Each trial lasted for 45 seconds.

[6] We drew new delays for all trials and participants. For bidirectional relationships, we assumed both directions shared the same delay.

During this time, components would activate if clicked on (constituting the interventions), or if caused by the activation of another component, with delay and probability governed by the true underlying network (Figure 7.5b). Graphically, a component activated through intervention turned yellow for 200ms and was marked by a "+" symbol. Components that were activated by another component turned yellow for 200ms but did not have the "+" symbol. Initially, all components were inactive and no connections were marked between them. Participants could activate components of their choice up to 6 times during each trial, and observe the resulting activations of the other components. The goal of the task was to identify the true connections. To incentivise the task and to get online structure judgments, we paid participants based on accuracy *during* the trials — i.e. based on whatever connections were marked correctly at a randomly chosen point during each trial. Thus, participants marked the connections during the trials by clicking in the grey boxes between each pair of components. This cycled through the options (*forward connection*, *bidirectional connection*, *backward connection* and *no connection*). The order in which clicks cycled through these components was counterbalanced. Every time the participant started to make changes to the connections a "confirm" button would appear in the middle of the screen and they would have to press this button to lock in their latest changes (Figure 7.5c).[7] A video clip of an example trial is available at `http://www.ucl.ac.uk/lagnado-lab/el/itv`.

Participants were trained about the delays in their condition through interaction with an an example device (always Loop-In-3, Figure 7.3) during the instructions. The device was paired with a picture showing the true relationships, and participants observed the components activating in sequence over 20 seconds with the delays reflecting the variability in their condition. To train participants on the response format, they had to mark the true links in the example and confirm them before they could move on. After the instructions, participants had to correctly answer comprehension check questions before they could proceed to the main task ensuring they understood: (1) the nature and number of interventions they could perform, (2) their incentive, (3) the 10% failure rate of the connections, and (4) the nature of the delays in their condition.

Participants then faced a practice problem (always Collider-3), and then the 12 devices in random order with randomly orientated and unlabelled components. When the 45 seconds ran out for each device, they were given feedback showing the true relationships and which of them they had correctly identified by the end of the trial (Figure 7.5d).

[7] This was done to distinguish participants' intended overall change from a sequence of singular changes as they cycled through different orientations and connections.

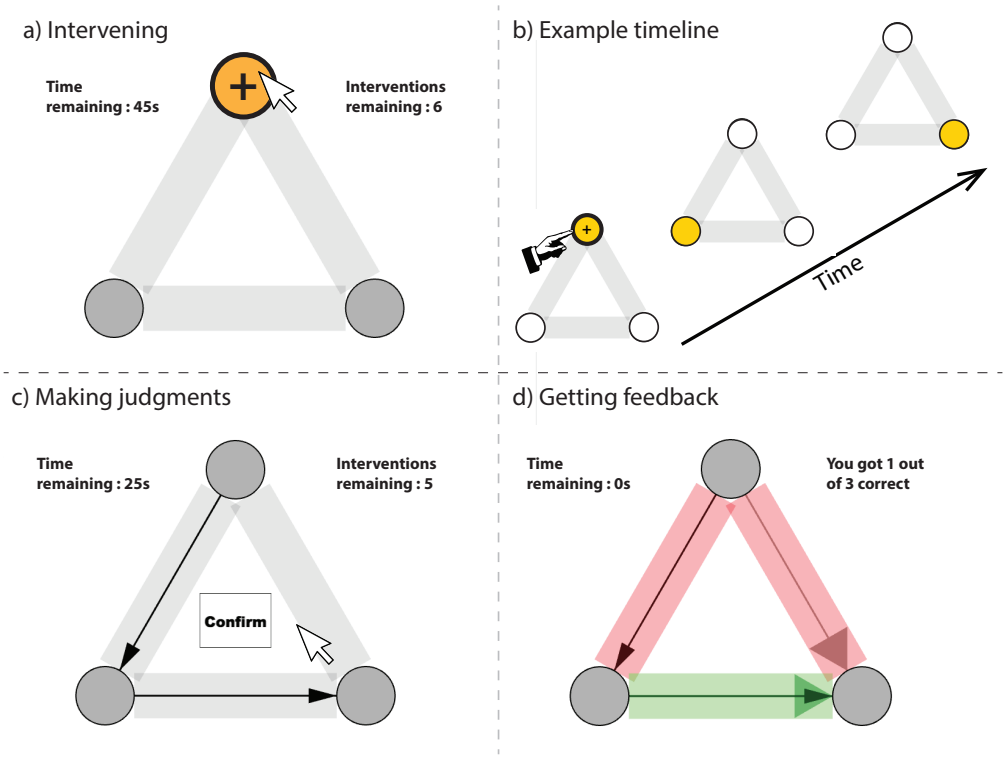


FIGURE 7.5: Procedure for Experiment 9. a) Up to 6 interventions could be performed by clicking on the components during the 45 second trial. b) This would lead to subsequent activations determined by causal connections and delays in the true model. c) Participants would also mark their beliefs about the structure during the trials by clicking on the edges. d) At the end of each trial they received feedback.

7.3.2 Results

Accuracy

Looking at participants' final judgments, we see they identified $0.75 \pm 0.19\%$, $0.63 \pm 0.16\%$ and $0.59 \pm 0.18\%$ of edges correctly in the *reliable*, *variable-within* and *variable-between* conditions respectively (see Figure 7.6). Average performance differed significantly by condition $F(2, 57) = 4.3, \eta^2 = 0.13, p = 0.02$. Post-hoc tests revealed that participants in the *reliable* condition identified significantly more edges correctly than those in the unreliable *variable-within* $t(38) = 2.1, p = 0.04$, and *variable-between* conditions $t(38) = 2.8, p = 0.006$. There was no significant difference between *variable-within* and *variable-between* $t(38) = 0.75, p = 0.5$. Participants confirmed judgments 1.6 ± 1.1 times per trial on average. Judgment time was not significantly related to accuracy, but final judgments were on average more accurate than initial judgments — 0.66 ± 0.31 compared to 0.62 ± 0.28 $t(719) = 5.8, p < .0001$. Accuracy by problem was highly correlated between conditions: *reliable–variable-within* = .91, *reliable–variable-between* = .86, *variable-within–variable-within* = .87. Only 5% of judgment updates decreased the

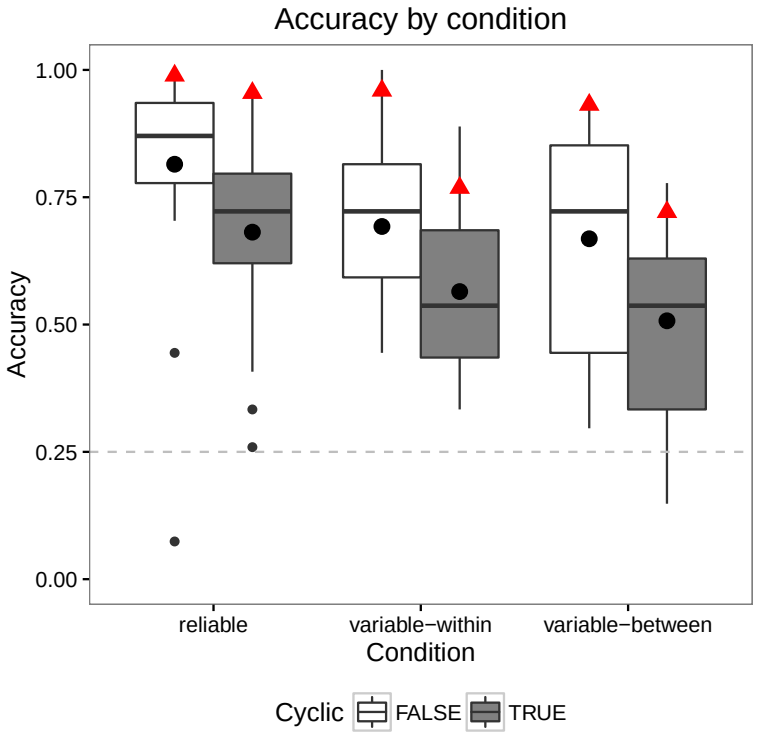


FIGURE 7.6: Proportion of edges identified correctly by condition. Boxplot shows medians and interquartile range for cyclic and noncyclic trials. Black circles = means. Red triangles = expected accuracy for guessing the maximum a posteriori model. Grey dotted line = chance.

number of connections marked, with 27% resulting in the same number as before, 68% increasing the number of connections.

Accuracy was also significantly lower for cyclic compared to acyclic models in all three conditions — *reliable* $t(19) = -3.8, p = 0.001$, *variable-within* $t(19) = -5.4, p < 0.001$ and *variable-between* $t(19) = -4.9, p < 0.001$. There was no interaction between the condition and cyclicity in predicting judgment accuracy. Inspecting Figure 7.7, we see that participants found the Loop-Out structures hardest to identify on average, struggling in particular to distinguish looping from output components. Accuracy was lowest for the *variable-between* condition, where simply observing the device cycle through many activations provided no new delay information.

In general, while participants' accuracy was considerably below the accuracy of an ideal learner that always judges the most likely a posteriori model (e.g. $\max P(M|\mathbf{d};\mathbf{c})$), the patterns of accuracy across condition and acyclic and cyclic devices are very similar (see Figure 7.6, comparing boxplot to red triangles).

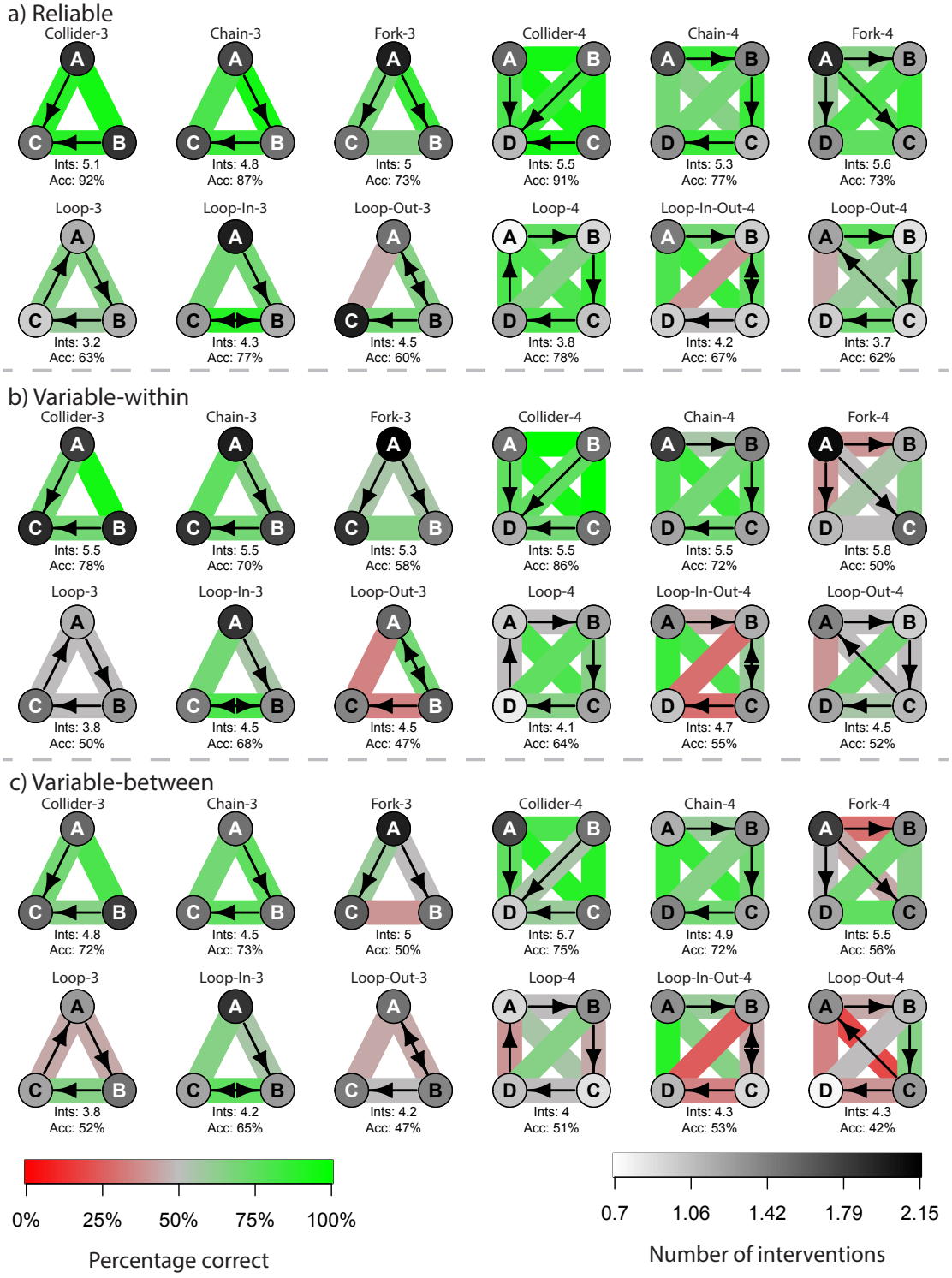


FIGURE 7.7: Intervention choice prevalence by component, and accuracy by edge in Experiment 9, in (a) *reliable*, (b) *variable-within* and (c) *variable-between* conditions. Average total number of interventions performed and accuracy written beneath each model.

Interventions

Quality Since, by definition, well chosen interventions lead to more information about the true model on average, we can assess the quality of participants' intervention choices indirectly by looking at the posterior uncertainty they afford by the end of each trial. Assuming normative inference and computing $H(M|\mathbf{d};\mathbf{c})$ (see Equation 3.8), we find that the quality of participants' interventions was a significant determinant of judgment accuracy over and above delay condition $t(56) = -3.3, \eta_p^2 = 0.16, p < .001$.

Spacing We hypothesised that spacing interventions out in time would be crucial to success. Accordingly, we measured the gaps between interventions for each participant. Participants waited 7.2 ± 2.8 seconds between interventions on average. Both the average length of the gaps participants left between interventions, and the regularity of these gaps were predictive of performance after accounting for delay condition $F(4, 55) = 12.2, \eta^2 = .47, p < .001$, with longer gaps $t(55) = 3.4, \eta_p^2 = .17, p < .001$ and less variability — as measured by the coefficient of variation $(\frac{\sigma}{\mu})$ for the inter-intervention intervals — $t(55) = -2.2, \eta_p^2 = .08, p = .03$ predictive of higher accuracy. Neither measure interacted with condition or one another in predicting accuracy.

A key question is whether well spread-out interventions were actually better at revealing the structures or merely a byproduct of generally successful causal learning. We can assess this by looking at the relationship between these measures and posterior uncertainties. If leaving bigger gaps and spacing interventions regularly are normatively sensible, we expect them to be negatively correlated with posterior uncertainty. We found that after accounting for condition, more widely spaced interventions were associated with lower posterior uncertainty $t(56) = -3.4, \eta_p^2 = 0.17, p = 0.001$ but the variability of these intervals was not $t(56) = 1.1, p = 0.2$.

Adaptation to cycles Participants performed fewer interventions on the cyclic devices (4.1 ± 1.1) compared to the acyclic ones (5.3 ± 0.8) $t(59) = 10, p < .001$ (see Figure 7.7). However, they still experienced far more activations in the cyclic systems (30.6 ± 11.3) compared to acyclic (4.5 ± 0.9) $t(59) = 17, p < .001$ due to the reciprocal relationships sustaining activations until one of the links failed. For cyclic devices, number of interventions performed was negatively related to accuracy $\beta = -.002, F(1, 58) = 12.7, \eta^2 = .18, p < .001$.

Positive testing We see evidence of a preference for positive testing, with participants performing 1.2 ± 0.5 times as many interventions per root compared to non-root component $t(59) = 3.9, p < .001$. This preference was associated with higher accuracy after accounting for condition $t(56) = 4, \eta_p^2 = 0.26, p < .001$ and did not interact with condition. However we note that, since a preference for interventions on root components depends on their successful identification, this relationship could be partly down to a relationship between successful identification of the root and accuracy. Consistent with our predictions, root preference was strongest for the root of the fork components in the *variable-within* condition with participants performing 2 ± 0.8 root interventions compared to 1.2 ± 0.6 on the other components for Fork-4.

Summary of results

In sum, we found that participants were better at identifying causal relations from intervention when delays were reliable, and the true structure was acyclic. Participants struggled particularly to identify the causal relations among loop-components and output-components. Successful participants tended to spread their interventions out more over the trial, distribute them more evenly and activate the root component(s) more often. The normative informational value of the chosen interventions was strongly predictive of final accuracy. Participants would frequently update their models by adding additional connections but rarely to remove connections.

7.4 Modelling the judgments

In the introduction we proposed four heuristic models of online causal structure induction. We now compare these to the patterns of judgments we observed from our participants. To do this we simulated belief trajectories $\mathbf{b}$ for all models, on all trials for all participants in all conditions starting from an unconnected model at $t = 0$. For the *Add most likely* and *Add more likely* models, we assumed knowledge of true μ, α and w_S as participants had been trained on these during the instructions. For the *Add more likely* heuristics, we had to set a threshold for how much more likely it had to be that an event was caused by a potential new cause than by the best existing cause in $b^{(i-1)}$ to justify its addition to $b^{(i)}$. We set this as a likelihood ratio of $\geq \frac{20}{1}$ in line with standard significance level of $p < .05$. Similarly, for the pruning step, we assumed the number of failures had to be significantly surprising at the $p < .05$ level to justify a connection's

removal from $b^{(i-1)}$ to $b^{(i)}$. After participants made a judgment we updated the $b^{(i-1)}$s of the models so that they matched the participant's current belief and updated them from there.[8] Having obtained simulated belief trajectories for all models, we took their current state at the moment of each participant judgment as their predictions. We then assessed their accuracy (e.g. the proportion of connections marked correctly) and their accordance rate with the participants (the proportion of connections marked the same as the participants'). For comparison, we also compared participants to a *Random* baseline that marked a new random causal structure for every judgment, and to a *Rational* learner that always selected the $\max P(M|\mathbf{d};\mathbf{c})$.

The results of these simulations are reported in Table 7.1. We see that all but one participant is better fit by one of the heuristics than by *Rational* or *Random* responding, with a large majority best described by the delay- and current-model-sensitive *Add more likely*, or the simplest *Add most recent* for both all judgments and restricted to the final judgments. *Add more likely* won overall in terms of correlation with participants and in terms of number of individuals best fit. There was little evidence for pruning with few individuals better fit, and no improvement in accuracy from inclusion of the additional pruning step. This may have been due to the relatively short trials leaving insufficient time for substantial model pruning, or potentially because our characterisation of pruning did not capture when and why participants did prune their structural beliefs. This also lines up with past studies in which participants would rarely remove connections after adding them (Lagnado & Sloman, 2006). Participants' accuracy (0.65 ± 0.19) was closest to that of the simplest heuristic *Add most recent*, with the other models somewhat more accurate than participants on average, and *Rational* judgment considerably higher again.

Table 7.1: Model comparison

Model	Accuracy (%)		Accordance (%)		N best (/60)	
	All	Final	All	Final	All	Final
Random	25	25	25	25	0	0
Add most recent	64	64	71	72	22	20
Add most likely	74	74	71	73	7	8
Add more likely	**78**	**80**	**74**	**77**	**26**	**30**
Add more likely + pruning	78	80	74	76	4	1
Rational	87	89	64	66	1	1

Note: All = for all judgments. Final = final judgments only. "Best fit" determined by the highest accordance rate (e.g. highest proportion matching connections across judgments).

[8] We also tried simulating the models without taking this step finding that they performed slightly better on average, but the distribution of individual best fits was very similar with the majority of participants according best with *Add most recent* or *Add more likely*.

7.5 Discussion

In Experiment 9, we found that people were able to use free selection of interventions to learn about the causal structure of devices whose dynamics took place in continuous time. As we expected, participants found cyclic structures harder to learn than acyclic structures, and this was reflected in the evidence, indicating that they were indeed normatively harder to identify in this setting, at least given the interventions that participants chose. This appeared to be due to it being hard to distinguish *output components* emanating from loops, from *loop components*. While this reveals a general difficulty for learning cyclic causal structures, it would also be possible to distinguish these if one could isolate parts of the device by blocking or turning off some of the components. Thus, in future work we would like to allow participants to perform a second kind of action, a block which prevents a component from activation until it is released, and see if participants can use this effectively to perform more controlled experiments.

There was no overall difference in accuracy between the *variable-within* and *variable-between* conditions, with similar drops in accuracy for both of these conditions relative to the reliable condition. These conditions also did not interact with other aspects of learning behaviour (e.g. distribution of interventions, accuracy on specific edges) as far as we could determine. This suggests that participants either did not make substantial use of within-connection variability in identifying the connections in the *variable-within* condition, or that any gains were offset by the greater predictability of the internally reliable delays in the *variable-between* condition.

As expected, we found that spacing out interventions over the trials improved the quality of evidence and ultimate judgment accuracy. Meanwhile spacing the interventions evenly was associated with better accuracy but not with generating more evidence from the perspective of normative learning. Thus even spacing of interventions might reflect a cognitive strategy for organising interactions with the device (e.g. into equal length semi-independent trials) rather than a normative requirement for generating evidence. Additionally, we found a preference for interventions on root components and a correlation between this and accuracy, but its prevalence did not differ in the variable conditions prohibiting any strong conclusions about the role of repeat testing and variability.

We found that we could explain participants' judgments best by assuming they added connections to a single evolving candidate hypothesis as the observed events, with some participants appearing to do so based on a simple most-recent priority heuristic, but more displaying evidence of sensitivity to delays (e.g. preferring an older event as a

cause when the most recent event was too recent to be plausible given the delays) and adding new links only when necessary (i.e. only if the current hypothesis really could not explain them well). This is consistent with the *Neurath's ship* perspective on causal learning as a process of local and incremental model construction.

7.5.1 Looking backwards or forwards?

More generally, the heuristics we proposed provide an interesting perspective on continuous-time causal model induction. Essentially, the heuristics worked by iterated *diagnostic* causal inference (Meder et al., 2014). Each time a new event occurred, the heuristics would look backward into the recent past for an explanation. The resulting perspective of causal model building is as backward-facing activity, even as its ultimate goal is to learn a forward model.

At first glance, a diagnostic focus seems to conflict with the idea of intervention choice, and active learning more generally, as inherently forward looking. Furthermore, research suggests that people find it easier to reason forward from causes to effects (Fernbach et al., 2011), and this gels with a simulation-based view of causal inference (e.g. Bramley, Gerstenberg, & Tenenbaum, 2016; Hamrick et al., 2016). Thus, it is interesting to consider how predictive and diagnostic inference might come together in this continuous-time setting. One possibility is that participants really did "look forwards" but in such a way that appears consistent with our heuristics. For example, learners might project something like a window of expectation forwards from each experienced event, automatically attributing subsequent events to the cause with the strongest current expectation (as in the blue density lines in Figure 7.2). It might be possible to construct heuristics that behave similarly to *Add most recent*, *Add most likely*, and *Add more likely* but do so based on projecting forward rather than diagnosing, so making expectation and violation the key tools for construction. This perspective chimes with Lagnado and Sloman's (2006) idea that events occurring shortly after an intervention are automatically attributed as effects.

Alternatively, it could be the case that learners used a mixture of backward and forward inference, perhaps attributing events following interventions as effects, but learning more diagnostically when observing sequences of non-interventional events (as would tend to occur in the cyclic models). Initially, with no idea where causal connections will be, interventions can only be used to search for variables that cause other variables to activate. But once model construction is under way, interventions will increasingly come

with expectations, so can be used to test specific or local hypotheses. In other words, we must diagnose enough candidate causal relationships to establish a forward model before we can start to use this to predict the future, act with expectations about results, and refine our causal models when these expectations are violated.

Chapter 8

Intervening in space

"*Ninety percent of life is just being there.*"

— ANDY CLARK paraphrasing WOODY ALLEN[1]

Chapters 3 to 5 focused on interventional causal learning in abstract scenarios, in which learners chose from a set of possible tests on a sequence of discrete trials. Chapter 7 opened the problem up to cases that lacked this discrete trial structure, in which interventions and resultant activity occurred in continuous time. However, the notion of intervention was still limited to activation of components of an idealised causal system. This final empirical chapter opens up the problem of active causal learning still further, looking at how people intervene and learn in physical "microworlds" with continuous spatiotemporal dynamics.

Nature's successful learners are embedded in the world they must learn to exploit, meaning they must construct their interventions from elemental physical actions. On this view, we can think of the little actions in everyday life as small experiments, ranging from the automatic (e.g. cocking one's head to better locate the origin of a sound), to the deliberate (lifting a suitcase to judge its weight; shaking a present to try and guess its contents; holding a pool cue to one eye, or spinning it, to gauge its straightness). A common element in these examples is that they seem to combine an intuitive understanding of physics with actions that exaggerate, or "bring into sharper relief" physical properties of interest. This implies that to have a better understanding of human active

[1]The percentage varies depending on the source. The original quotation is probably "80 percent of success is showing up" (Safire, 1989).

causal learning, we can look at how people use their ability to act directly in the physical world, to test and discover causal properties.

In this chapter we begin the process of exploring this naturalistic type of learning. We do this by looking at how people learn about physical properties, such as local magnet-like forces and object masses when interacting with simulated physical "worlds". We start by briefly surveying the literatures on active learning and intuitive physics learning, then describe an experiment that contrasts passive learners with active and yoked learners. Finally, we looked at the information generated by observing compared to actively intervening on the worlds, and begin to categorise the types of actions, or "natural experiments", that active participants performed. We find that active learners use their ability to control the worlds to generate information differentially, creating situations that are highly informative about target properties while minimising confounding "noise" information about irrelevant properties.

8.1 Active learning in richer domains

Human active learning has largely been studied in abstract scenarios where the space of possible actions is limited, such as the Wason card selection task (Oaksford & Chater, 1994), category rule learning (Gureckis & Markant, 2009) and games like "Guess Who" (Nelson et al., 2014) and "Battleships" (Markant & Gureckis, 2010). A related line of research has explored active causal learning, where the available actions are more overtly physical, involving interventions on idealised causal systems (Bramley, Lagnado, & Speekenbrink, 2015; Coenen et al., 2015; Lagnado & Sloman, 2004). Since many causal structures cannot be distinguished by covariational data alone (Spirtes et al., 1993; Steyvers et al., 2003), the concept of intervention captures a key aspect of real-world active learning that goes beyond simply asking the right questions. Learners can make use of their ability to act causally in the world to gain first-hand experience of cause–effect relationships. The large majority of causal learning research has focused on CBNs (Pearl, 2000) where time and space are abstracted away, and actions are limited to idealised interventions (Bramley, Dayan, et al., 2017; Bramley, Lagnado, & Speekenbrink, 2015; Bramley, Nelson, Speekenbrink, Crupi, & Lagnado, 2014; Coenen et al., 2015; Lagnado & Sloman, 2002, 2004, 2006; Steyvers et al., 2003). In general, these studies found that people intervene in ways that provide more information than passive observation or random intervening, but that their choices also tend to be more stereotyped and repetitive than those prescribed by models of optimal active selection.

This has led to proposals that active learners' choices are better understood as boundedly rational (Bramley, Dayan, et al., 2017; Bramley, Dayan, & Lagnado, 2015; Bramley, Lagnado, & Speekenbrink, 2015) meaning they tailor their actions to their own limited learning capacities, testing only a subset of the possible hypotheses at any given time, and limiting the computational cost of inference by controlling for confounding factors.

If learners' actions are heavily tailored to their idiosyncratic learning trajectories, we expect the evidence they generate to be less useful for other learners, with different idiosyncrasies, observing their choices (Markant & Gureckis, 2014). This view is broadly (Lagnado & Sloman, 2004; Sobel & Kushnir, 2006), but not always (McCormack et al., 2016), supported by experiments that have included yoked conditions, where one participant observes the tests performed by another. Intuitively, the divergence between information that is in principle available and what participants can actually learn, will be much larger in more complex and naturalistic situations, where only a fraction of the total evidence can plausibly be attended to.

While emerging research has begun to explain how people can learn intuitive physical theories and use these to infer the physical properties of objects (Battaglia, Pascanu, Lai, Rezende, et al., 2016; Chang, Ullman, Torralba, & Tenenbaum, 2016; Ullman et al., 2014; Wu, Yildirim, Lim, Freeman, & Tenenbaum, 2015), to the best of our knowledge, no one has yet explored how active learning shapes this process in humans.

8.1.1 Intuitive physics

In recent years, research into people's intuitive understanding of physics has experienced a revival. This is partly due to the ease with which we can design physically realistic displays thanks to available software packages with physics engines. While early research into intuitive physics had focused on documenting how people's understanding of some aspects of physics, such as ballistic and curvilinear motion, is sometimes systematically biased (e.g. McCloskey et al., 1980), more recent research has demonstrated how some of these biases may be explained if we assume that (1) our physical understanding is approximately Newtonian, and (2) we are often fundamentally uncertain about some important aspects of the physical scene (e.g., the masses of the objects involved in a collision, Sanborn et al., 2013).

Battaglia et al. (2013) have argued that people's understanding of physics is best understood in analogy to a physics engine used to produce physically realistic scenes. The

idea is that people have something like a physics simulator in their mind that they can use to approximately predict what will happen in the future (Smith & Vul, 2013), reason about what happened in the past (Smith & Vul, 2014), or simulate what would have happened if some aspect of the situation had been different (Gerstenberg et al., 2015). From the causal model perspective, we can think of this as a set of particularly rich constraints on the form of a causal model — e.g. where relationships between entities are given functional forms that embody the equations of motion, conservation of energy and so on.[2] Nevertheless, the results of these experiments are consistent with the view that people have a rich intuitive theory of physics that supports approximately accurate mental simulations of key aspects of physical scenes. However, these experiments do not address the question of how we get there – how do people acquire their intuitive physical theories?

Intuitive theories can be expressed as probabilistic programs (Gerstenberg & Tenenbaum, to appear; Goodman et al., 2008, 2015). Such programs can contain both logical and stochastic functions meaning that they are capable of generating a distribution of actualisations. For example we might express a theory of attraction as a very general version of Newton's universal law of gravitation Force $= G\frac{m_1 m_2}{d^2}$ in which the strength of the attraction is defined by the masses of the relevant objects m_1 and m_2, the distance between them d and (gravitational) constant G. However, we can treat m_1, m_2, d, G, or even the $\frac{1}{x^2}$ functional relationship with d, as random variables, each with their own probability densities. The result is a program, or theory, which can be used to generate a wide range exact forms of physical attraction, but that still expresses constraints, ruling out many forms and making some of the remainder more plausible than others.

Probabilistic program induction is a thorny problem, but one where human-like performance has been demonstrated (Lake et al., 2015; Lerer, Gross, & Fergus, 2016) by sophisticated Bayesian machinery embodying principles of causality and compositionality. Ullman et al. (2014) explored human intuitive physics learning by studying how people learn about different latent physical properties of 2D "microworlds" similar to the one shown in Figure 8.1. The worlds were bounded by solid walls and contained a number of coloured pucks with differing weights, surfaces with differing levels of friction, as well as local (magnet-like) forces between pucks and a global (gravity-like) force pulling all the pucks in a particular direction. The properties of the worlds (the number

[2] Of course the idea that we have a physics simulator in our head raises the crucial question of how the brain can implement this. While we do not focus on this here, note that there are a number of emerging proposals based on deep learning (e.g. Battaglia et al., 2016; Chang et al., 2016; Ullman et al., 2014; Wu et al., 2015).

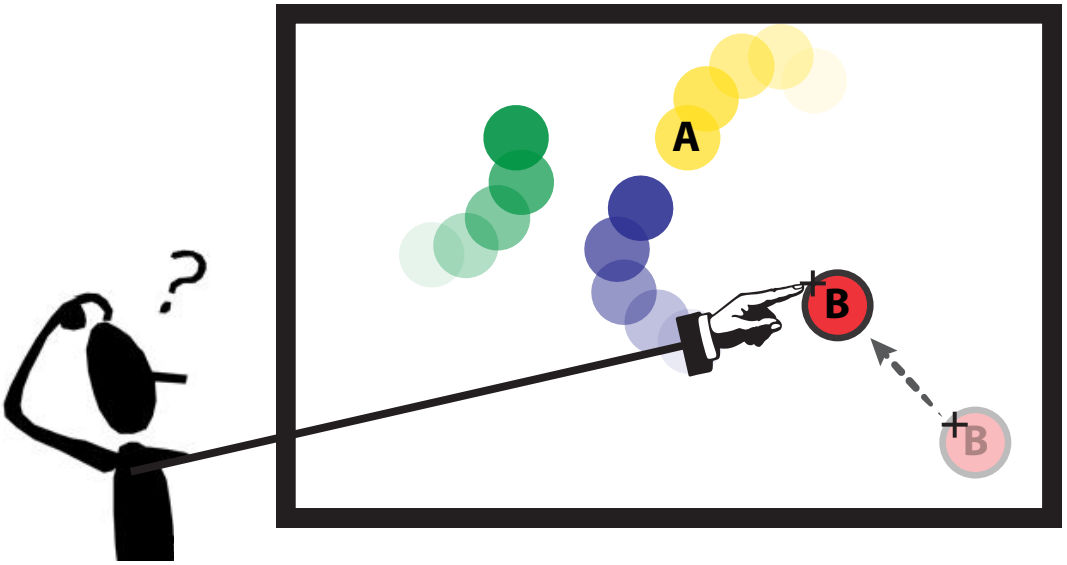


FIGURE 8.1: Schematic display of "microworlds" in Experiment 10. Four pucks are moving around colliding and affecting one another with local (magnet-like) forces. The learner is also dragging the puck labelled "B" (by left-hold-clicking on it and moving the computer mouse).

and nature of the pucks, friction patches and forces) were generated from an underlying probabilistic program capable of generating around 14,000 distinct worlds. Participants would watch and then replay a 5 second clip from each of the generated worlds. In each clip, the pucks bounced around, attracting and repelling each other, being slowed down by the friction, and being pulled by the global force. Participants then answered a series of questions about each world's properties. Participants were able to detect different levels of mass and friction on average, but individual judgments were noisy. They identified the correct global force around 70% of the time and were much better at detecting local attraction (82%) than repulsion (53%). Ullman et al. found that divergence was matched by an asymmetry in the evidence: pucks that repelled one another would rarely spend long enough close together to exhibit strong repulsion, while attracting pucks would rapidly approach one another and stick together offering stronger evidence of the latent force.

Ullman et al. modelled participants' judgments by assuming a mixture of an Ideal Observer model (IO) and a Simulation and Summary Statistics model (SSS). The IO compares the observed objects' trajectories to simulations of expected trajectories under the different possible worlds. The model assumes a certain amount of perceptual uncertainty (e.g. about the pucks' exact locations and velocities) and uses this to calculate likelihoods of the observed data under different possible world settings. The SSS compared statistics about each clip such as the pucks' average positions, velocities and pairwise

distances, to the summary statistics of repeated simulations under the different possible worlds. For instance, objects in worlds with a global force towards south tend to be closer to the southern wall of the world. A mixture model that combined both IO and SSS had a .81 correlation with participants' judgments. Individually, the SSS did a better job than the IO on predicting all but global force judgments.

In the current work, we build on these results, exploring how people interact with physical microworlds and their learning of the different physical properties. The SSS approach is less well suited to the active learning setting, because it is unclear how interacting with the worlds will affect their summary statistics. In general, we would expect perturbing the world to make these a less reliable guide to the worlds' properties. For instance, if you move pucks to the southern wall of the world, then their average location is no longer a good guide to the global force. The IO perspective is more promising, providing a way of quantifying the evidence provided by interactions with the worlds relative to evidence received without intervening. We can also use the IO perspective to distinguish *what* learners' actions provide the most information about, and more generally, how acting in the worlds reshapes the evidence available for learning.

8.2 Experiment 10: Intervening in space and time

The final experiment of this thesis uses a task adapted from the setup used by Ullman et al. (2014). However, rather than preselecting scenes to show participants, we generated the simulations on the fly. This allowed us to include an *active* condition in which participants could exert control over the scenes and alter how they played out, as well as a *passive* condition in which participants merely observed the world and a *yoked* condition in which participants observed the actions of an active participant. There are many ways in which we might allow active learners to interact with the worlds. Bramley et al. (2016) piloted two active learning setups that differed in the extent to which participants had fine-grained control. In an *"active punch"* condition, participants controlled a fist with which they could roughly knock other objects around, mimicking the clumsy actions of a baby yet to develop fine motor skills. In an *"active grab"* condition, learners could use the mouse to grab the pucks and drag them around, staging more precisely orchestrated interventions. Participants found the fist hard to control, leading to chaotic interventions and low accuracy. Therefore, in this experiment we focus on the *active grab* condition.

We also focus on identification of two target properties: local pair-wise forces, and object masses. Because active testing is particularly valuable when competing causal explanations cannot be resolved by observational evidence only, we generated confounded evidence by including two distractor pucks along with two target pucks and drawing local forces randomly out of attract/none/repel for all pairs of target and distractor objects. This means that it is important to isolate the target pucks from the distractor pucks to get clear information about the target pairwise force. Rather than including global forces, or friction which were easily identified by passive learners in Ullman et al. (2014), we varied the relative mass of the two target objects, a property which participants had found more difficult to infer, and whose identification we hypothesised would benefit from curated comparisons and interactions between pucks.

We hypothesised that *active* participants would outperform *passive* participants, and that *yoked* participants would inherit some, but not all of this advantage.

8.2.1 Methods

Participants

Sixty-four participants were recruited from Amazon Mechanical Turk (39 male, $M \pm SD$ age 33.6±10.2). Participants were paid at a rate of $6 per hour, plus performance-related bonuses ($0.61 ± .17).

Conditions

Participants were assigned to one of three learning conditions, *passive* ($N = 24$), *active* ($N = 20$), *yoked* ($N = 20$):

1. **Passive** Participants observed the microworlds unfold without being able to interact. If, in rare cases, everything came to a standstill, objects' velocities and locations were refreshed.
2. **Active** Participants could grab pucks and drag them around with the mouse. Grabbed pucks retained their properties (i.e. mass and local forces and location and momentum) but became strongly attracted to the position of the mouse.

When released, they would continue on their current trajectory but no longer be attracted to the mouse.[3]

3. **Yoked** In this condition, participants watched replays of the interactions of an active participant.

The first 44 participants were randomly assigned to either the *passive* (24) or the *active* (20) learning condition, and the final 20 were *yoked* 1-to-1 with the 20 active participants.

Worlds

Each participant watched or interacted with 9 microworlds, consisting of all combinations of target force in *attract*, *repel* and *none* and target masses in $[1, 1]$kg, $[1, 2]$kg and $[2, 1]$kg (see Table 8.1). The five other pairwise forces were drawn uniformly from the three possibilities for each participant on each trial. This resulted in an overall hypothesis space $\mathcal{M}$ of 2187 possible worlds (e.g. all 3^7 combinations of target and distractor local forces and the possible target masses) but a smaller judgment space containing the 9 combinations of target mass and target force.

Table 8.1: Experiment design

World	1	2	3	4	5	6	7	8	9
Target force	A	A	A	N	N	N	R	R	R
Target 1 mass	1	1	2	1	1	2	1	1	2
Target 2 mass	1	2	1	1	2	1	1	2	1

Note: A = attract, N = none, R = repel; masses are in kg.

Materials and Procedure

The experiment was programmed in javascript using a port of the Box2D physics game engine (demos of all three conditions and and replays of all *active* participants are available at `https://www.ucl.ac.uk/lagnado-lab/apl`). The microworlds were displayed in a 600 by 400 pixel frame, with 1 m in the world corresponding to 100 pixels on the screen. Each world was bounded by solid walls with high elasticity (98% of energy retained per collision) – and contained four pucks of random colours[4] that were different on each trial. All four pucks had radius of .25 meters and elasticity of 98%. The two

[3] We opted for strong attraction rather than simply copying the position of the mouse because this allowed the controlled object to interact reciprocally with the other objects in collisions rather than behaving as if it was infinitely heavier than the other objects.

[4] These were chosen to be equally spaced around the HSV colour wheel (Smith, 1978) with a random starting point each time.

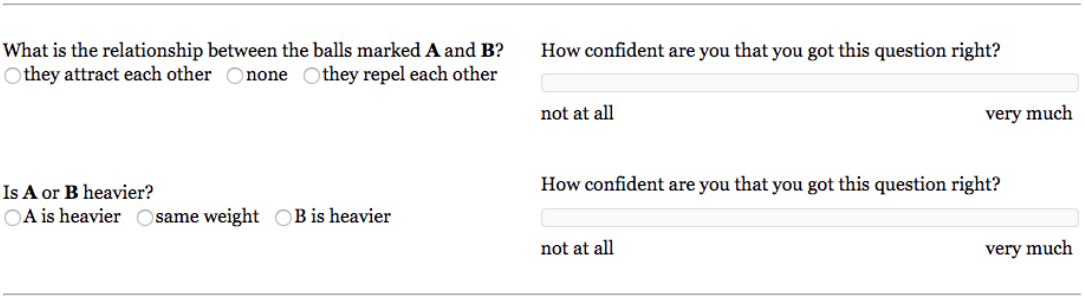


FIGURE 8.2: Eliciting judgments in Experiment 10. Participants answered two forced choice questions and set confidence sliders.

target pucks were labelled with new letters on each trial (e.g. "A" and "B" on trial one, "C" and "D" on trial two, cf. Figure 8.1) while the distractor pucks were unlabelled. This was done to minimise transfer effects and confusion between the objects in the different trials which had been an issue in Bramley et al.'s (2016) pilot. The distractor pucks were all 1 kg but one of the target pucks could weigh 2 kg as in Table 8.1.

Each world also had up to 6 distinct local forces, one between the target pucks, and one for every other combination of target and non-target puck. Each of these could either be attractive ($3\,\mathrm{m/s^2}$), repulsive ($-3\,\mathrm{m/s^2}$), or no force.[5] The pucks' initial positions were random but non-overlapping, with initial velocities in the x and y direction drawn from $\mathrm{Unif}(-10, 10)\,\mathrm{m/s}$. If all pucks' velocities fell below $.15\,\mathrm{m/s}$, the simulation froze and the window went black for 500 ms before the positions and velocities of the pucks were redrawn.[6] Each world was simulated for 45 seconds at 60 frames per second, leading to 2700 frames of evidence per trial. Complete specification of the settings of the Box2D simulator and a demo of the experiment are available in Appendix C.

At the end of each trial, our two test questions appeared in counterbalanced order (see Figure 8.2). The questions were forced-choice but were paired with confidence sliders for a fine-grained measure of participants' judgments. To ensure that participants were motivated to answer the questions as well as they could we paid a bonus for each correct response.

Participants first completed instructions relevant to their condition, answered comprehension check questions, and then faced two practice trials followed by the nine test trials. Practice trials were always worlds 1 and 5. The randomly drawn distractor forces, puck colours and labels differed between the practice and test instances. The two

[5] Local forces scaled with the inverse squared distance between the objects in line with Newton's universal law of gravitation. Thus, the current local force L exerted on object o_1 by object o_2 (and the reverse) was given by $\frac{3}{d^2}$.

[6] This happened 1.0 ± 0.83 times per 45 second trial on average.

test questions appeared below the world when the time was up. At the end of the experiment, participants received feedback about how many of the test questions they got right, and were paid a 5¢ bonus for each correct answer. The experiment took 19.0 ± 7.3 minutes on average.

For yoked participants, the cursor of the participant to whom they were yoked (hereafter the yoker) was shown with a large "+" symbol whenever it was within the world, and any objects grabbed by the yoker were indicated as in the active condition with a thick black border. Counterbalancing and puck colouration from the active condition was shared by each yoker–yokee pair.

8.2.2 Results

Accuracy

Participants answered 53%, 66% and 54% of questions correctly in the *passive*, *active* and *yoked* conditions respectively (see Figure 8.3). Average performance differed significantly by condition $F(2, 61) = 3.8, \eta^2 = .12, p = .03$. Post-hoc tests revealed that *active* participants answered significantly more questions correctly than passive participants $t(42) = 2.5, p = 0.02$, and their yoked counterparts $t(19) = 2.9, p = 0.02$, with no difference between passive and yoked participants $t(42) = .2, p = 0.83$. Only 4 yoked participants outperformed their active counterparts (Figure 8.3, dotted lines), with a further 3 answering the same number of questions correctly. Active participants' performance was predictive of their yoked counterparts' $F(1, 18) = 5.6, \eta^2 = .24, p = .03$.

Confidence judgments differed by condition $F(2, 61) = 5.3, \eta^2 = .15, p = .007$, with active participants significantly more confident on average than passive $t(42) = 2.8, p = .006$ or yoked participants $t(38) = 2.9, p = .006$. Confidence was positively correlated with accuracy $F(1, 62) = 10.6, \eta^2 = .15, p = .002$ but did not interact with condition.

Masses versus relationships

Across conditions, participants were worse at inferring masses than forces $t(63) = -4.8, p < .0001$ and reported lower confidence in mass judgments ($66 \pm 25\%$) compared to force judgments ($74 \pm 25\%$) $t(63) = -4.2, p < .0001$. Again, participants were less accurate in correctly identifying when there was *no force* between the target pucks (56%) than *repulsion* (70%) or *attraction* (78%), with a main effect of question type

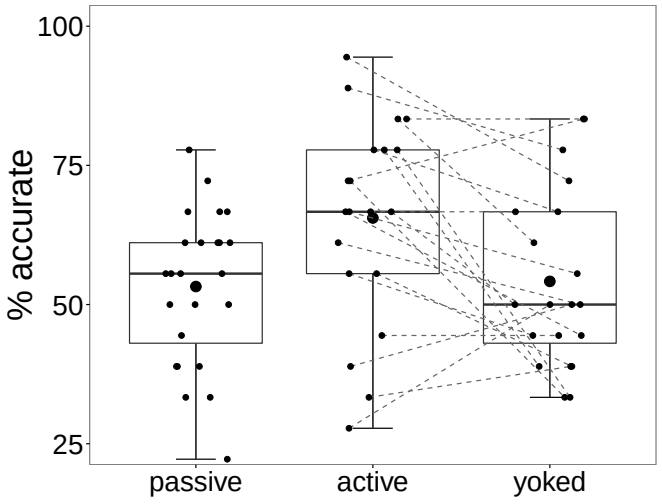


FIGURE 8.3: Performance by condition in Experiment 10. Boxplot shows medians and interquartile range. Large dots indicate condition means. Small dots indicating individual participants are jittered for visibility. Dotted lines connect active participants with matched yoked participants.

$F(2, 189) = 7.7, p < .0001$ and significant improvements going from *no force* to *attraction* $t(126) = 3.9, p = .0001$ and *repulsion* $t(126) = 2.4, p = .017$. Force type additionally interacted with condition $F(6, 183) = 3.0, p < .0001$. Dummy contrasts with *no force* and *passive* as controls revealed active participants were significantly better at identifying *repel* than passive participants $t(42) = 3.2, p < .0001$ and there was a marginal improvement for yoked participants as well $t(42) = 1.9, p < .058$ (see Figure 8.4). There was no significant relationship between accuracy on the local force question and the number of distractor forces.

Effects of control

Active participants experienced slightly fewer between-puck collisions than passive participants, 59 ± 14 compared to $65 \pm 9, t(42) = 2.0, p = 0.056$. However, they experienced significantly more collisions between the two target pucks 15.0 ± 8.1, compared to $9.8 \pm 4.4, t(42) = 2.7, p = 0.01$. $13.2 \pm 7.8\%$ collisions in the active condition took place while one of the two target objects was being controlled by the participant. Time spent controlling objects was also positively related to final performance for active and yoked participants $F(1, 38) = 4.8, \eta^2 = 11, p = 0.04$.

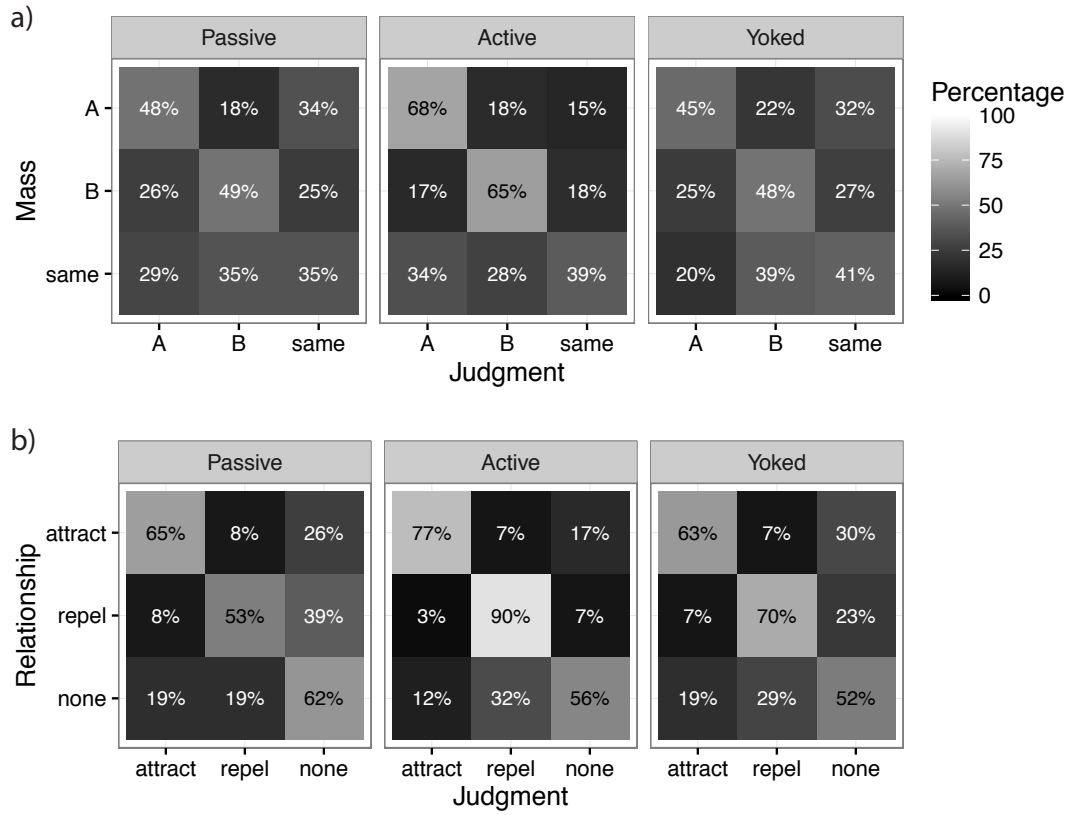


FIGURE 8.4: a) Confusion matrices for mass question (a) and relationship question (b).

8.2.3 Interim discussion

The markedly better performance in identifying repel forces by the active participants is striking, considering that *passive* participants and those in Ullman et al. (2014) struggled with this property. This is consistent with the idea that active participants were able to force the repellent pucks closer together and thus gain more experience of these forces in action. In general, we saw a marked improvement in performance when participants could control the scenes themselves. Thus, key questions are *what* the active participants were doing, and *how* their actions helped them identify the worlds' properties. In the next section we begin to explore this both quantitatively — by measuring the evidence generated throughout each trial — and qualitatively — by categorising the different testing strategies the active participants came up with.

8.2.4 Measuring information

Similar to Ullman et al.'s (2014) *Ideal Observer* model, we used simulations to compute likelihoods for different worlds given the true trajectories of the objects $\mathbf{d}$ in each trial.

We simulated each trial under all 2187 possible world settings and tracked how much they diverged from what actually happened. Because of the dynamic nature of the physical interactions, simulations with only slightly different properties would quickly diverge from the same starting point, becoming completely unrelated after a few seconds. Thus, to get a balanced measure of the evidence available throughout the trials, we snapped the simulated objects back to their true locations and velocities every 10 frames (6 times per second) before allowing them to start to diverge again.[7] We converted these divergences into likelihood scores assigning a probability of observing the actual object trajectories $\mathbf{d}$ given the potential world m_i, the interventions $\mathbf{c}$ (in the active and yoked conditions), and some Gaussian perceptual noise. We calculated these as

$$p(\mathbf{d}|m_i;\mathbf{c}) = \prod_{t=1}^{T} e^{-\epsilon(s_i^t - d^t)^2}, \tag{8.1}$$

for each object, where d^t is its location at time t, s_i^t its simulated location if m_i is true, and ϵ is a scaling parameter capturing perceptual uncertainty about the objects' true locations and velocities (see Figure 8.5a for a visualisation of this procedure).[8] By computing these likelihood scores for all possible models and on each trial, multiplying by a uniform prior $P(M)$, we computed a posterior over models $P(M|\mathbf{d},\epsilon;\mathbf{c})$ and associated posterior uncertainty $H(M|\mathbf{d};\mathbf{c})$ using Equation 3.8. We also calculated posterior uncertainty relative to the target mass and relationship questions by marginalising over the other properties. The posterior uncertainty depends on what interactions occur during a clip. For example, objects must pass close together for their trajectories to be strongly affected by any local forces and must collide with one another (or be dragged by the mouse) for their trajectories to be strongly dependent on their masses. Thus, this measure provides a way of assessing whether active learners moved the objects in ways that generated extra information relative to *passive* learners.

We find that overall posterior uncertainties did not differ on average for passive compared to active participants $t(29) = 1.5, p = 0.13$, nor did posterior uncertainty about

[7] Like Ullman et al. (2014), we experimented with several "snap back" windows, including 5 and 20 frames, finding similar results.

[8] Note that we do not assess the scaling parameter for perceptual noise ϵ, so these measures should be thought of as a guide to the *relative* rather than *absolute* evidence in the trials. We assumed an $\epsilon = 10$. $\epsilon = \infty$ corresponds to perfect knowledge of the objects' locations and velocities, which, combined with perfect knowledge of the physics engine, rules out all but the true world within a few frames. $\epsilon = 0$ would assign equal likelihoods for all worlds.

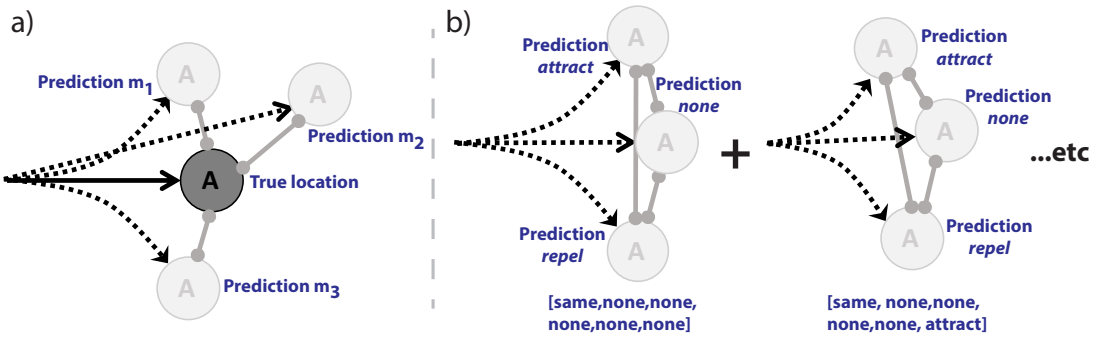


FIGURE 8.5: a) Calculating model likelihoods based on divergence of simulated trajectories. Black line and opaque puck shows the true trajectory of one of the objects in a scene. Dotted lines and transparent pucks show simulated trajectories under different assumptions about the world. Note that a simulation with the correct assumptions about the world would lie on top of the black line. We based the likelihood for each hypothesis m on the distances between its simulation and the true trajectory (grey lines). b) Measuring *predictive divergence* due to target force, averaging over all other properties. Rather than measuring distance between simulated and actual trajectories as in (a), we measured the distance between simulations that differed on only one property at a time, matching all other properties (shown at the bottom in square brackets), and averaging across all possible other properties (indicated by "+" and "...etc"). Note we still snapped these back to the true trajectories every ten frames.

mass $t(29) = 0.8, p = 0.4$, but that active participants generated significantly more information on average about the target force $t(29) = 3.8, p < 0.001$.[9] Inspecting Figure 8.6, we see that mass uncertainty was generally very low (meaning the ideal observer should normally be very certain about the mass question) significantly lower than force uncertainty $t(60) = -7.1, p < .001$. Mass uncertainty was also skewed for active participants in particular, with mainly low but a few very high values. Thus we also compared log uncertainties, now finding a significant difference between active and passive in terms of log mass uncertainty $t(29) = 3.9, 0 < 0.001$, and no difference in results for logs of overall and force uncertainty — $t(29) = 1.6, p = 0.13$ and $t(29) = 3.9, p < .001$ respectively.

Intervention quality

As well as using simulation to estimate what participants managed to learn by the end of the trials, it is also interesting to look *within* the trials to see which periods better revealed which properties. For this we created a measure we call *predictive divergence* (PD). This captures the extent that the current action in a trial *depends* on the value of a given property. We calculate PD for each world-property using the simulated object trajectories, splitting them according to their value on each property (e.g. whether A

[9] Unfortunately velocity information was not stored for 13 participants meaning we could not use their clip data for computing these measures. Thus, the following comparisons are based on the remaining 18 passive and 13 active participants (and the 13 corresponding yoked participants where appropriate).

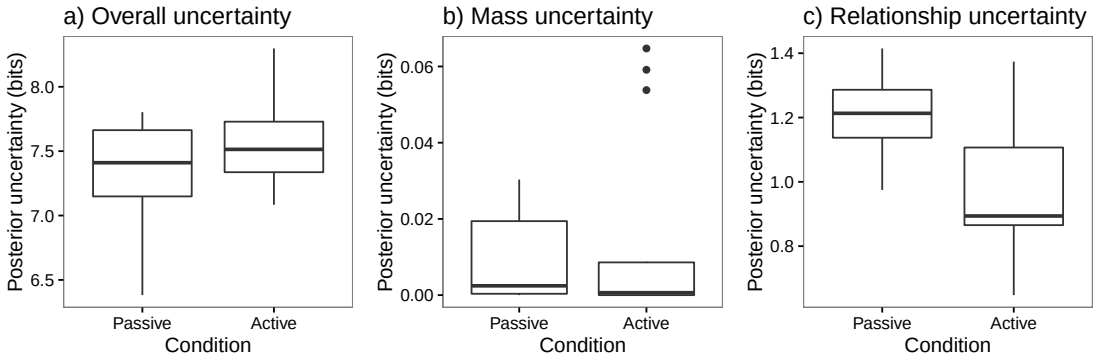


FIGURE 8.6: Remaining uncertainty for active and passive participants: a) With respect to the true world. b) With respect to the target masses. c) With respect to the target force.

attracts, has *no effect* on, or *repels* B), measuring the degree that these predictions diverge and averaging this over possible settings of the other properties. The result is a measure of how strongly each property is being revealed at every point in every trial (see Appendix C for equations). To a first approximation, maximising this measure is a good objective for planning interventions. Interventions with large expected PDs correspond to questions in which there is no mistaking the answer; one can expect very different things to happen depending on the truth about the target property, irrespective of the other properties of the world. We consider three variants. PD_{mass} measures the current predictive divergence depending on the target objects' masses. PD_{force} does the same for the target force. As a baseline, we also consider PD_{any} which is the average predictive divergence for any (target or distractor) property of the world. As a concrete example, if target objects are close together and at rest they will have high PD_{force}. This is because they will move toward each other if they attract, stay still if there is no force, and move away from each other if they repel, leading to strongly differing predictions about their trajectories depending on this property.

Accordingly, we find that active participants have much higher average PD_{mass} $t(29) = 5.6, p < 0.001$ and PD_{force} $t(29) = 4.6, p < 0.001$ than passive participants but do not differ in terms of PD_{any} $t(29) = 0.4, p = 0.7$ (see Figure 8.7a). Additionally, looking within active learners' trials, we compare periods of active control to periods of passive observation. We find that periods of control average only a moderate increase in PD_{any} $t(11) = 3.2, p = 0.009$, but a large increase in PD_{mass} $t(11) = 10.2, p < .001$ and PD_{force} $t(11) = 5.1, p < .001$ (see Figure 8.7b).

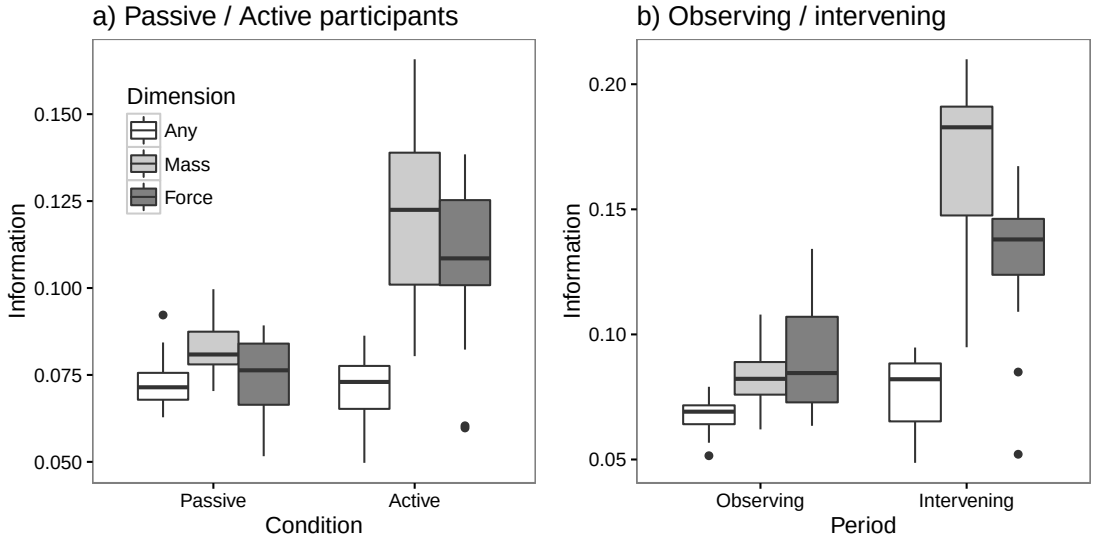


FIGURE 8.7: Comparison of (a) predictive divergence for passive and active participants, and (b) between periods of observation and intervention for active participants.

8.2.5 Natural experiments

Clearly, active participants are acting in ways that are effective at revealing the target properties over and above the irrelevant distractor properties. Understanding exactly how they came up with the actions they did is a large project. However, as a first pass, we viewed the replays and identified number of potential strategies. We describe these here, providing schematic figures (8.8a–f)and links to replays exemplifying them:

(a) **Deconfounding** Even though participants mainly manipulated the target pucks, they also sometimes manipulated the distractor pucks. Many of these manipulations involved moving the distractor pucks out of the way and leaving them at rest in a far corner (Figure 8.8a, `http://www.ucl.ac.uk/lagnado-lab/el/it/de`).

(b) **Encroaching** Participants grabbed one target puck and brought it toward the other target puck. This simple strategy allowed participants to infer whether and how the two pucks affected one another. In some cases, participants towed one attracting puck with the other, or pushed a repulsive puck around with the other providing a strong and extended demonstration of the force between the pucks (Figure 8.8b, `http://www.ucl.ac.uk/lagnado-lab/el/it/encroach`).

(c) **Launching** Participants grabbed one of the target pucks and "threw it" against the other target puck. This intervention helps to figure out whether one of the targets is heavier than the other (Figure 8.8c, `http://www.ucl.ac.uk/lagnado-lab/el/it/launch`).

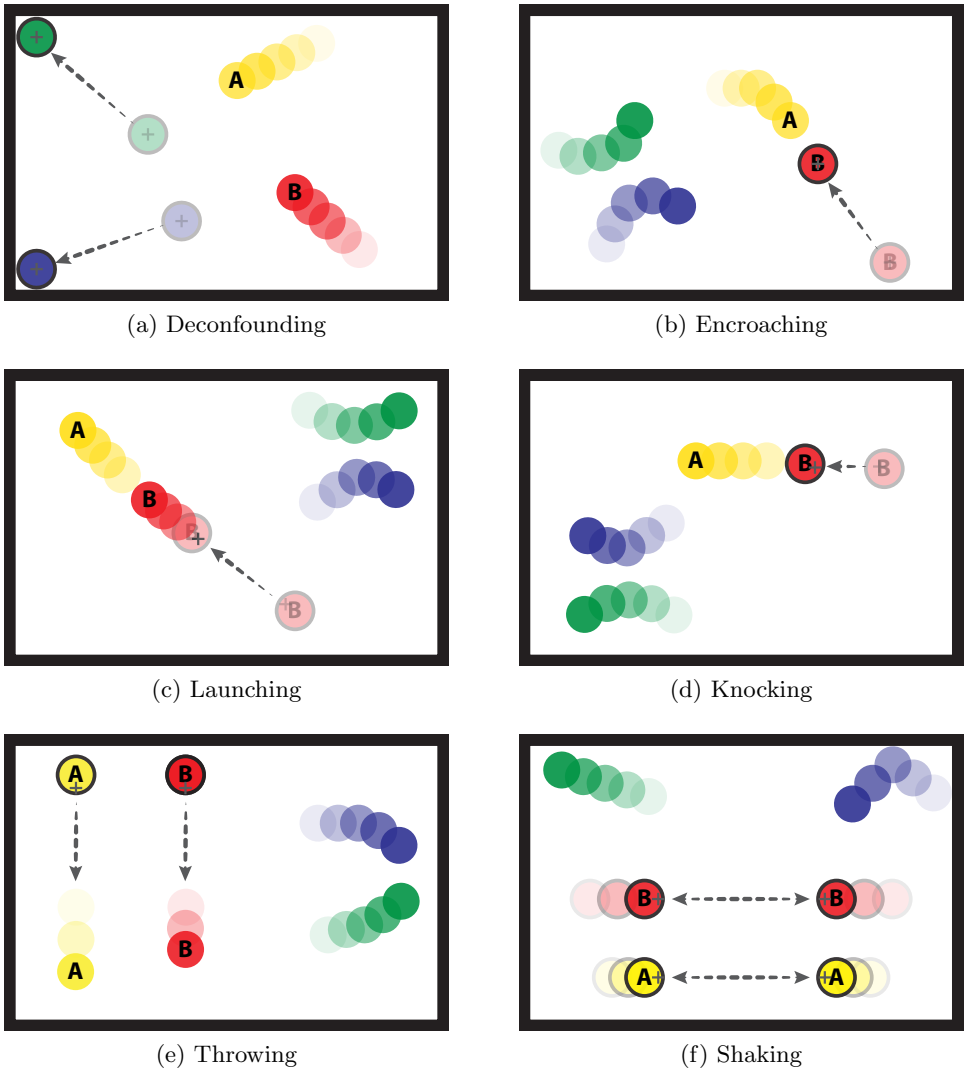
(d) **Knocking** Similar to *launching*, participants grabbed one of the target pucks and knocked it against the other (without letting it go). This intervention also reveals information about the mass of each object (Figure 8.8d, `http://www.ucl.ac.uk/lagnado-lab/el/it/knock`).

(e) **Throwing** Participants grabbed a target puck and then threw it, explicitly avoiding collision with any of the other pucks. By exerting an identical force when throwing each target puck, the results of the intervention help to figure out the mass of each object (Figure 8.8e, `http://www.ucl.ac.uk/lagnado-lab/el/it/throw`).

(f) **Shaking** Some participants discovered an effective strategy for comparing the mass of the two target objects. By rapidly shaking each in turn (moving the mouse from side to side) it was possible to see that the heavier object reacted more sluggishly. Its greater momentum takes longer to be counteracted by its attraction to the mouse location (Figure 8.8f, `http://www.ucl.ac.uk/lagnado-lab/el/it/shake`).

Some of these strategies have easy-to-measure hallmarks. For instance, in line with *encroaching* (Figure 8.8b), we see evidence that participants in the active condition identified the local forces by bringing the two target pucks close to each other. The lower the average distance between two target objects for an *active* participant, the better they did on the force question $\beta = -.3, F(1, 18) = 8.0, \eta^2 = .3, p = .001$ but this had no relationship with accuracy on the mass question $p = .87$. Conversely, in line with the *shaking* strategy (Figure 8.8f), participants who moved the controlled object around faster did better on the mass question $\beta = 25, F(1, 18) = 15, \eta^2 = .45, p < 0.001$, but controlled object speed had no relationship with accuracy on the force question $p = .67$. Yoked participants did not inherit these differences, with no significant relationships between performance on either question and average distance between targets or controlled-object speed.

We also explored the relationship between these potential strategies and the PD measures. Figure 8.9 shows the information generated in a trial that we categorised as a *shaking* strategy. From around halfway through the trial onward, the participant shakes the target balls one at a time, strongly revealing their relative mass. In line with our intuition, we see this creates large spikes in $\mathrm{PD}_{\mathrm{mass}}$ relative to $\mathrm{PD}_{\mathrm{force}}$ and $\mathrm{PD}_{\mathrm{any}}$. Figure 8.10 shows the $\mathrm{PD}_{\mathrm{mass}}$ and $\mathrm{PD}_{\mathrm{force}}$ generated throughout a trial categorised as both *encroaching* (bringing target objects close together) and *disambiguating* (moving distractor objects out of the way), again particularly in the second half of the trial, we see spikes in $\mathrm{PD}_{\mathrm{force}}$ relative to $\mathrm{PD}_{\mathrm{any}}$, but also many spikes in $\mathrm{PD}_{\mathrm{mass}}$ (which was generally higher than $\mathrm{PD}_{\mathrm{any}}$).



(a) Deconfounding (b) Encroaching

(c) Launching (d) Knocking

(e) Throwing (f) Shaking

FIGURE 8.8: Examples of different interventions participants performed in the *active* condition.

8.3 Discussion

In sum, we found a clear benefit for active over passive learning in this experiment. In particular, active participants gathered more evidence about repulsion by bringing target objects closer together and moving distractor objects out of the way. A number of them also gathered information about masses by shaking the target objects back and forth. The quality of the control exerted by the active participants was an important determinant of the quality of the final evidence available to the yoked participants. However, the substantial drop-off from active to yoked accuracy was consistent with the idea that first-hand knowledge of *what* was being tested (e.g. relationship or mass), *when* and *how*, was likely to be crucial for learning successfully. Another factor might

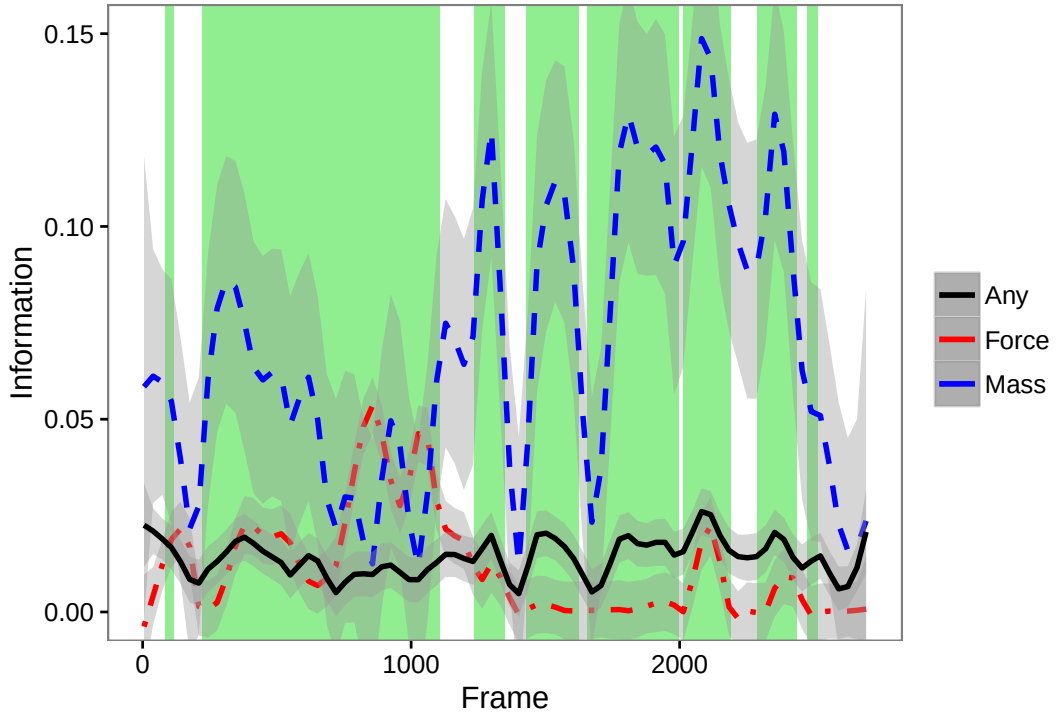


FIGURE 8.9: Timeline for an active trial exemplifying *shaking*. Periods of control shown by green shading. Smoothed PD_{mass} (blue dashed), PD_{force} (red dot-dashed), PD_{any} (black full) estimates with light grey confidence intervals. Watch the clip: `http://www.ucl.ac.uk/lagnado-lab/el/it/shake`.

have been that active participants were able to look ahead at the crucial locations in the scenes where diagnostic interactions were expected to occur. Yoked participants lacked the ability to foresee what will happen.

There are a number of ways we might model these differences between active and yoked performers' experiences. In our initial analysis, we treated all objects' locations and velocities as equally uncertain. However, it is plausible that active learners have a better idea about the locations of objects while controlling them since they can incorporate direct motor feedback from their mouse or finger on the track-pad (e.g. Körding & Wolpert, 2004). We could model this by giving active learners a smaller perceptual uncertainty parameter for objects under control. This would mean that active learners receive stronger evidence from events involving the controlled object. Additionally, learners' attention is certainly limited relative to the action in the scenes. Thus, we might model learners' attention as a focal window. Active learners could then use their knowledge of planned action to move their window, for instance toward regions they expect to be informative in virtue of their interventions. Yoked learners lack this foresight and hence may have often be looking elsewhere when something informative happens.

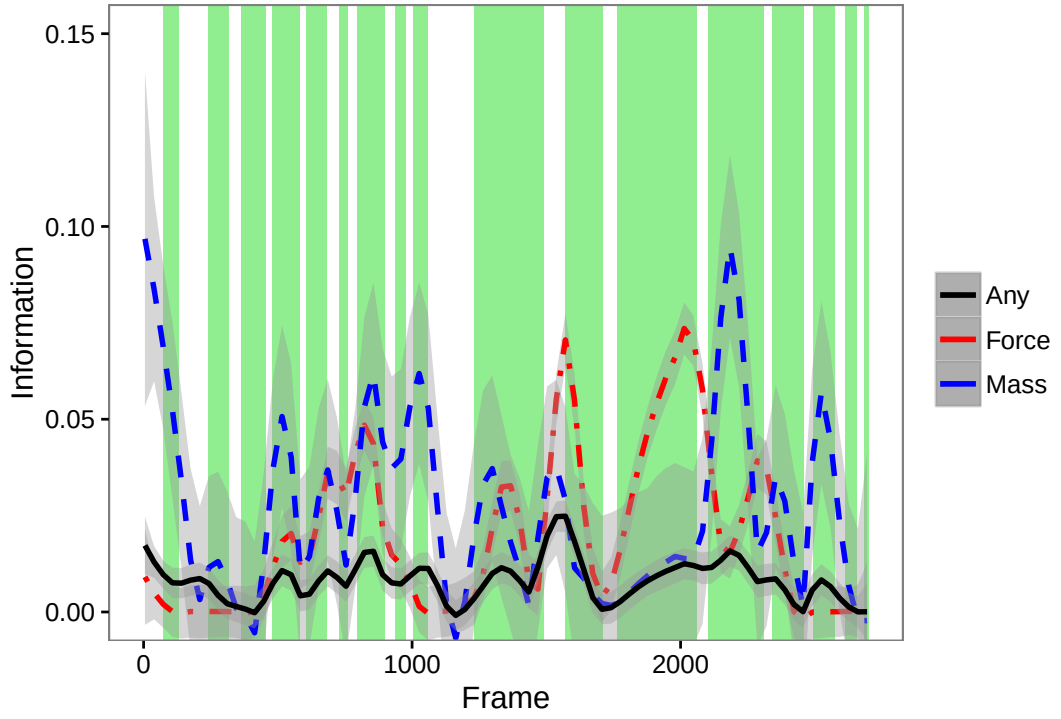


FIGURE 8.10: Timeline for active trial exemplifying *deconfounding*. As in Figure 8.9. Watch the clip: `http://www.ucl.ac.uk/lagnado-lab/el/it/de`.

An interesting disconnect between our measures and participants' judgments was that the measures suggested there was more information available about masses than forces, while participants found the mass question much harder. There are several possible explanations for this. One possibility is that participants were also uncertain about other aspects of the world. For simplicity we assumed perfect knowledge of the worlds' fixed properties (e.g the elasticity of the objects, the friction, the strength of the attractive force of the mouse on controlled objects, the laws of the simulated physics). It could be that incorporating uncertainty about these other properties makes the model likelihoods and predicted divergences less sensitive to differences in mass. Another possibility is that types of divergence caused by the local forces were easier for people to spot. The local forces created *qualitative* differences in the paths of objects (e.g. making objects veer toward or away from one another rather than continuing in a straight line) while the masses affected things more *quantitatively* (e.g. affecting the degree of veering or the angle of exit collisions). It is plausible that the perceptual system is better tuned to distinguishing curves from straight lines than comparing angles, especially as these veering movements are hallmarks of causal influences. Indeed, research has suggested that people are poor at predicting the quantitative consequences of collisions (Sanborn et al., 2013; White, 2006a).

Encroaching and *shaking* permitted simple indirect measures, and accordingly, we found shakers doing better on mass questions and encroachers doing better on relationship questions. Additionally, we found our predictive divergence measures lined up with our intuitions about the value of *shaking* and *disambiguating* strategies in particular. The other experiments' timeline signatures were more subtle, depending more on what happens after control is released (e.g. throwing) and we plan to explore these in future work. One planned step is to have independent coders watch the replays and code instances of proposed strategies. This will give us a more reliable idea of strategy identifiability as well as a basis from which to model them. Another is to repeat the experiment, tasking participants to infer only mass *or* force property. This will allow us to look at whether our information measures, or yoked participants can identify what the active participant is learning about, and assess the extent that matched versus mismatched goals affect the responses of yoked participants.

In the longer term, we would like to explore the idea that people have intuitive theories of active learning, which they can use to rapidly generate informative actions in familiar domains. Work on this problem has begun in machine learning. Denil et al. (2017) use deep neural networks combined with reinforcement learning to train an algorithm that can generate physics experiments in a virtual environment (see also Chang et al., 2016). Similarly, Agrawal, Nair, Abbeel, Malik, and Levine (2016) train a robotic arm to actively learn about the physical properties of real objects by poking and prodding them. These projects found success through explicitly or implicitly encoding a hierarchical model of the action space — loosely speaking, meaning that action planning can be done in a top-down way; so choosing whether to *shake* or *knock* before generating (or simulating possible) motor-realisations. Our experiment supports the idea that humans are intuitive natural scientists, quickly finding ways of revealing causal properties in a physical (but novel) setting. It seems likely that future work in this area will benefit from combining insights from machine learning with close analysis of human behaviour.

8.4 Conclusions

This chapter began the process of exploring active learning of intuitive physical theories. We found that active learners spontaneously generated informal experiments that allowed them to identify the relevant properties more effectively than passive controls, but that their success was not shared by yoked observers. We demonstrated that active learners generated considerably more information about target properties and discussed

the types of interventional strategies, or "Natural experiments", participants performed and outlined steps for future work.

Chapter 9

General discussion

> "*Thirty years ago, we used to ask: Can a computer simulate all processes of logic? The answer was 'yes', but the question was surely wrong. We should have asked: Can logic simulate all sequences of cause and effect? And the answer would have been 'no'.*" — GREGORY BATESON

> "*Know thyself*"
> — ANCIENT GREEK APHORISM

9.1 Causal theory change reconsidered

My initial rational analysis of causal structure learning cast it as a problem of Bayesian inference over a vast hypothesis space of possible models. However, we quickly saw that explicit maintenance of probabilistic beliefs in such a large space would be impossible for any plausibly bounded learner. As a solution, I conceptualised causal learning as an incremental process of "mental tinkering", where one gradually "tries out" small adaptations to a single global working model or theory with the goal of improving its fit, or at minimum, accommodating the latest evidence. Approximating intractable probabilistic inference has received a lot of attention in machine learning (e.g. Bishop, 2006). Accordingly, the mental tinkering idea readily found formal footing in a model (*Neurath's ship*) that cast theory change as a semi-stochastic local search in model space, veering between a broad Markovian search (Gibbs sampling) or a race to the nearest peak (hill-climbing). Relying on local search meant the current structure could act as an anchor allowing the learner to benefit from evidence learned in the past but now

forgotten, and providing a rationale for conservatism and anchoring effects in cognition (Edwards, 1968; Hogarth & Einhorn, 1992; Lieder et al., 2012; Petrov & Anderson, 2005). I noted at the end of Chapter 5 that my approach is not limited to CBN structure learning, but that variants might explain human success at learning all sorts of complex representations. For instance, in Chapters 6 and 7 patterns of causal learning from point events in continuous-time were consistent with the idea that learners considered new evidence through the lens of the current working hypothesis, which they would adapt as required to explain unexpected or surprising new evidence. Chapter 8 did not explicitly assess the proposal that people adapt their beliefs about physical causality in a local and incremental manner. However, á propos the previous chapters, it seems plausible that learners fixed their assumptions about non-focal aspects of the worlds when predicting the future, and so updated their beliefs about the world, one aspect at a time.

In general this anti-foundationalist perspective lines up with common intuitions about the role of thinking cognition. We expect that complex problems will take time to percolate, and we often engage in actively rethinking a familiar issue with hope of discovering a better or deeper understanding. It can take a lifetime, or even multiple generations, to discover a simpler solution or more powerful explanation. These inherently sequential and stochastic aspects of cognition, which we often term "thinking", are difficult to accommodate from the computational level Bayesian perspective, where all the possibilities must be laid out in advance. However, they seem natural from the perspective of a cognitive agent as engaged in stochastic search and optimization in a latent large and multi-modal theory space.

9.1.1 Getting started

While *Neurath's ship* was the best fitting model of participants in Chapter 5 and the related *Add more likely* model accorded best with participants in Chapter 7, there was also substantial empirical evidence for the use of heuristics. For instance, many participants in Experiments 1–4 would attribute all effects directly to their latest intervention (*simple endorsement*) and in Experiments 5–9 people often behaved as if the temporal order of events was equivalent to their causal order. In one sense this is puzzling. Complete reliance on such rules would seem to prohibit incremental refinement. One reason for their prevalence in our data could be that our experiments focused on short learning periods and demanded causal judgments early and repeatedly. This forces learners to

come up with ways of getting started — i.e. generating an initial hypothesis that they can lean-on and progressively refine. It could be that simple endorsement and temporal order heuristics are ways of getting started rather than sufficient strategies in their own right, meaning that their prevalence in the current experiments is in part due to the focus on short learning trials and unfamiliar settings.

9.1.2 Navigating infinite hypothesis spaces

The *Neurath's ship* characterisation of incremental causal model construction works within a predefined — albeit large — hypothesis space. However, human learning and development may be better characterised as a process of *discovery*, in which the final model is not in any meaningful sense, contained in the original formulation of the problem. Thus to understand how new hypotheses are discovered, we must rethink what kinds of theory modifications are available to the learner. One element of this could be to operationalise model changes as application of *production rules* (e.g. Goodman et al., 2008). The idea here is that the learner has a small set of primitives that they can chain together to create an infinite number of new hypotheses. In a simple causal context these might simply be variables and edges. If the learner can not only add and remove edges but also create variables — perhaps by dividing existing ones (Buchanan et al., 2010) or positing latent ones (Lucas, Holstein, & Kemp, 2014) — then they can, in principle, create any causal model of arbitrary complexity. On this view, theory change could proceed through a search that stochastically adds and removes these primitives, allowing the model grow in complexity as the data demands (Aldous, 1985; Antoniak, 1974; Ghahramani & Griffiths, 2005; Pitman, 2002).

Taking this one step further, developing theories might often require the positing of new primitives — e.g. when the existing ones are simply not expressive enough to accommodate the evidence. We see this in physics where new entities — e.g. strings, bosons, black holes — are often initially proposed as placeholders to fill gaps in existing theories. In graphical model terms, we might imagine positing a new kind of edge such as a bidirectional link $\leftrightarrow$ as a new primitive to explain dynamics we cannot accommodate with just $\leftarrow$ and $\rightarrow$. As a real-world example, the modern concept of magnetism was once a place holder for curious behaviour of iron ore (Verschuur, 1996), and germ theory was born from John Snow's observation of geographical patterns of cholera cases clustered around water pumps (Snow, 1855). Thus an interesting direction for developing a theory of cognitive theory change might be to formalise additional moves such as the addition

and removal of variables, and generation of new primitives. In terms of the metaphor, we can imagine a *Neurath's ship* that is capable of sailing into uncharted waters (via production rules), and being jerry rigged with new parts made from flotsam (unexplained data) encountered on the way. Given formal flesh, such a theory could begin to explain how complex cognition can emerge from almost nothing through a few simple moves and an extended interaction with the causal world.

9.1.3 Vertical structure

An important aspect of both cognitive representation and active learning that I have not focused on explicitly, is its vertical structure. The fact that our participants learned successfully across a range of domains from the highly abstract (Chapters 3–5) to the richly physical (Chapter 8) confirms the intuition that humans are capable of reasoning about causality across levels of abstraction, and both with and without domain specific knowledge. I noted in the introduction that *hierarchical Bayesian models* have been used to express theories of situation specificity — in which existing causal knowledge can be rapidly recruited in the form of rich priors in familiar domains (e.g. Tenenbaum et al., 2011). Furthermore, I noted that *recursive Bayesian networks* (e.g. Casini et al., 2011) can model the process of reductive explanation — explaining phenomena in terms of their parts — and abstraction — generalising about the aggregate behaviour of parts. An interesting and open question is whether we can express the *Neurath's ship* style ideas about local and incremental belief change in such a "three dimensional" theory space. The thought is that lower level posits are nested theoretically or ontologically in higher level posits, such that changes at one level can impact downwards (affecting what lower-level parts are needed to explain a high level phenomenon) and upwards (affecting what higher-level phenomena emerge from abstraction) as well as horizontally (affecting the roles of other local relationships). While these sorts of ideas have been developed in philosophy of science (e.g. Strevens, 2008), they are yet to be explored in detail in cognition.

Constructing a thesis

A prominent recent example of complex theory construction for me was writing this thesis. At no point did I select this thesis wholesale from among all possible theses, but rather constructed it phrase by phrase. The finished product is the result of a (seemingly endless) chain of small additions, edits and subtractions, during which I

leant increasingly on the other sections and chapters for support. But I would also shift my focus upwards and downwards between different levels of organisation. Some days I focused at a high level, reordering the chapters or thinking about the overall message. These high level changes would tend to have unforeseen consequences that I would gradually discover later when working on the content of individual chapters. For instance it would no longer make sense to foreshadow a certain result, or would be too late to introduce a certain idea. Likewise, small local changes could later turn out to have consequences at the level of high level organisation. Additionally, in line with the challenges of learning a causal model, the hardest part of writing the thesis turned out to be getting started. The final draft of the introduction is only distantly related to the first draft, and I tended to sketch an initial draft of each chapter rapidly — e.g. Chapter 7 as the shortest route between Chapters 6 and 8 — without addressing nuance or detail. Once I had something to build on, I would then spend a much longer time focusing more narrowly, editing a paragraph at a time supported by the surrounding work-in-progress.

9.2 Intervention reconsidered

9.2.1 Optimal self teaching

This thesis was primarily about active learning. The rational analysis perspective on intervention selection revealed it to be an even more intractable problem than the theory building it supports. Thus I found success in modelling people's behaviour under the assumption that, to the extent that they chose their interventions in order to resolve uncertainty, they did so relative to much narrower questions than the global "what is the true causal model?". Studying active learning alongside structure inference, helped make sense of phenomena like confirmation driven, local and repetitive testing as normatively sensible ways of dealing with a limited ability to learn from evidence. The goal of active learning for a bounded cognitive agent is not, as often presumed, the generation of as much information as possible, but rather a process of self-teaching, turning the world into a classroom in which one's limited learning system can flourish and understanding can grow. Just as a teacher who simply presents the hardest material without ensuring the students are following would be a bad teacher; an active learner who generates informative situations without consideration of their own learning trajectory would be a bad active learner. Thus, one of the key messages of this thesis is that we should rethink

optimal active learning as *optimal self-teaching*, and accept that it should be shaped as much by the limitations of the student, as by dimensions of the world.

The divergence between generating information and successful learning increased in breadth throughout the thesis. In the three-variable discrete-trial learning problems there was a large amount of overlap between the actions of an ideal learner and those of participants. But as we looked at domains of increasing richness and complexity — i.e. the real time interventions of Chapter 7 and ballistic "experiments" in the physical microworlds — it was easy to get too much of a good thing. A noiseless idealised learner with perfect knowledge of the physical rules could identify the target properties after a few frames of activity in the physical microworlds of Experiment 10. However, for real learners with limited attention and uncertainty about the physical rules, it was hard enough to identify these properties within 45 seconds, requiring curated interactions between the objects for high chance of success. Thus we found that successful learners' actions tended to be those that selectively magnified evidence about target properties while minimising the influence of causal distractors. In the same way, we would expect a good teacher to focus on presenting his students with clear cases and unambiguous demonstrations of basic principles, safe from the complexities and confounds that complicate learning outside the classroom.

9.2.2 Interventions that explore versus interventions that test

We might divide the types of interventions people performed into two camps: (1) Open-ended activity-generating interventions, that participants would tend to perform earlier during learning and (2) controlled tests that recognisably focused on learning about a particular element of the problem. This suggests a dual role for interventions that has been suggested informally before in other contexts (e.g. Ruggeri & Lombrozo, 2014). Early on, interventions are good for hunting causes (Figure 9.1a, Cartwright, 2007). By systematically trying things out — as we might characterise children's play (e.g. Bruner et al., 1976) — we can start to discover putative effects and "get started" on model construction. As our model becomes more developed it starts to make predictions. By comparing these predictions under a few local counterfactuals we start to be able to predict what would happen, for instance, if a particular hypothesised causal relationship does exist versus what would happen if it does not exist (Raiffa, 1974). This predictive ability provides a basis for selecting interventions pro-actively, to resolve local uncertainty. We can start to imagine what tests create strongly diverging predictions

depending on some manageable local question, against the backdrop of the rest of larger theory (e.g. Figure 9.1b). By choosing interventions whose potential evidential impact we have already pre-emptively imagined, we naturally restrict our reach to the performance of tests whose outcomes we will be able to interpret. Indeed, on the preposterior analysis view, we have already done the work by the time we choose our intervention. Upon seeing the actual outcome, all that remains is to travel down the associated arm of our prediction tree.

This notion of intervention might also shed light on confirmation bias (Klayman, 1995; Nickerson, 1998). Distinct from (but related to) confirmatory evidence gathering, confirmation bias is a tendency to interpret evidence as favouring one's hypothesis more than it really should. From the preposterior analysis perspective, this might stem from classifying actual outcomes as one of a narrow set of predictions. Where there is an approximate or near-fit between a prediction and the outcome, this may be taken as confirmation even though a detailed assessment would reveal the inadequacy of the question. This echoes Lakatos (1976), who talks of degenerate research programs surviving through their "internal momentum" — e.g. through preoccupation with a narrow set of tests and results to the neglect of the larger picture. On the plus side, the Neurath's ship perspective demonstrates that, when balanced appropriately, such a narrow focus need not prohibit the gradual improvement of one's theory.

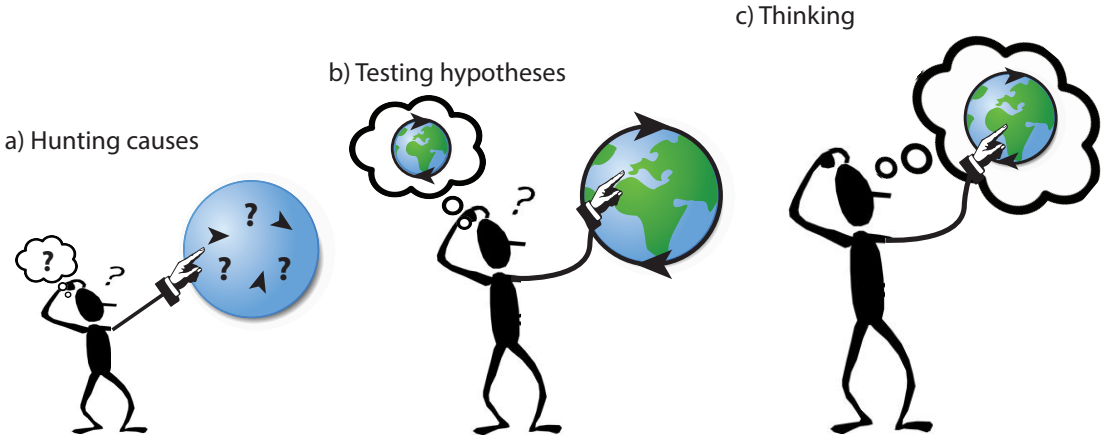


FIGURE 9.1: Three stages of active causal cognition. a) Early in development interventions help us discover causes. b) Once we have a model, they let us refine it by testing hypotheses. c) Once the model is rich enough, we can interrogate it through mental intervention to think, problem solve and control.

9.2.3 Intuitive theories of intervention — learning to actively learn

How can it be that almost all our participants were able to come up with interventions that were robustly more informative than chance despite facing abstract complex learning problems in which online calculation would be computationally expensive? The answer is probably that they didn't. Our participants did not have to derive each intervention from first principles. Rather, they could rely on a mature intuitive theory of action and its relationship with causality, and more proximally on the discovery of effective, repeatable strategies over the course of the problems in an experiment. For example, if turning X_1 "on" tells you what X_1 affects, then turning X_2 "on" will tell you what X_2 affects. If you have controlled for one confound, you can do the same with another. If confirmatory testing worked in one learning scenario it may work again for a similar one. If shaking is informative about the mass of one object then it will be informative about the mass of another. In general, we do not need to reinvent the concept of an experiment every time we want to learn something. We can rely on learning a *theory of intervention*, similar to those posited to explain other aspects of cognition (e.g. Gerstenberg & Tenenbaum, to appear; Lake et al., 2015; Tenenbaum et al., 2011). From this perspective, we can use our knowledge of good strategies for testing things to learn quickly in familiar domains. All we need to do is work out what kind of learning problem we are facing and try what worked in the past. Thus, a fruitful avenue for future work would be to apply insights about learning and representing intuitive theories to the domain of active learning. This could provide insight into how people are able to come up with creative and revealing learning strategies on the fly in complex situations like the microworlds in Experiment 10. Indeed, it is here, where the participants presumably had much stronger domain knowledge, that we saw the most creative and distinctive active learning. It is hard to see how the kinds of behaviours we saw from participants could be constructed from scratch on the spot. More plausibly, a rich hierarchical theory of action allowed them to work top down, using high level learning goals like "get a clear look at the target force in action" (Chen & Klahr, 1999; Klahr et al., 1993) to generate candidate strategies (e.g. ways of getting the distractor pucks out of the way), which in turn would guide the generation of motor realisations.

9.2.4 From intervention to adaptive control

As I depicted in Figure 2.1 (Chapter 2), causal learning and representation is really, ultimately, in service of a more general goal: exploiting the world. I did not focus on the

problem of balancing exploration — the use of interventions to gather information – and exploitation — acting to gather rewards (Macready & Wolpert, 1998; Schulz et al., 2016; Shanks et al., 2002; Steyvers et al., 2009), and was deliberately agnostic about the precise relationship between utility, information and reward. However, striking the right balance between learning and exploiting the causal world is clearly an important question. The jury is still out in neuroscience on whether there are truly separable value signals for information and reward (Daw & Doya, 2006; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006). Thus, it may be that we can understand cognition better by unifying these goals in a single model. One way of achieving this is through the notion of Bayes-adaptive control (Bellman, 2015; Feldbaum, 1960; Klenske & Hennig, 2016). Under an adaptive control analysis, both learning a model and exploiting it are dual aspects of a single decision problem that can be expressed as a partially observed Markov decision process (Puterman, 2009). Specifically, we can model control as the task of maximising a reward in a state space where the transition rules (e.g. the causal structure) is initially unknown. The possible trajectories in such a Bayes-adaptive Markov decision process (or BAMDP, Guez, Silver, & Dayan, 2012) encompass all possible future beliefs — i.e. all ways that the world might turn out to work. Solving this decision tree tells you what action to take now in order to maximise long-run expected future returns. Unsurprisingly, the exact solutions to non-trivial BAMDPs are intractable, requiring consideration of a generally unmanageable number of possible futures. However, recent work has discovered a range of powerful approximations — based on sophisticated sampling and aggressive deep-tree-search techniques (e.g. Guez, 2015; Guez et al., 2012; Guez, Silver, & Dayan, 2013). Thus, a future direction for studying causal cognition and active learning in general might be to work within this unified formulation of the problem. From here we might start to make sense of intuitively common aspects of thinking, such as playing out scenarios far beyond predictability, as achieving something akin to the deep tree search heuristics that allow for a good approximation of optimal control in a complex and uncertain causal world.

9.3 Turning control inward

A continuing disconnect between artificial and human intelligence is task flexibility. While machines now regularly outperform humans on individual pattern matching or complex control tasks (Guo, Singh, Lee, Lewis, & Wang, 2014; LeCun, Bengio, & Hinton, 2015; Mnih et al., 2015), they typically have a narrow domain of competence, and

flounder in areas that require more flexibility such as problem solving and creativity (Lake, Ullman, Tenenbaum, & Gershman, to appear). A potential explanation for this continued divergence is the idea that active learning plays a special role in cognition which is lacking in passively trained AI systems. In developing an intuitive theory of intervention rich enough to guide interaction with the world, humans might also be developing the skills they need to begin to exert control *internally*: interrogating their internal models, and taking control over their own cognition (Figure 9.1c). A number of aspects of cognition including decision making and hypothesis generation have an active character. For instance, decision by sampling theory (Sanborn & Chater, 2016; Stewart, 2009; Stewart et al., 2006) suggests that we interrogate our memory, generating samples which we use to rank and choose between options. Similarly, diagnostic hypothesis generation has been characterised as an active probing and generating of examples from memory (e.g. Navarro & Perfors, 2011). This idea is a natural partner for causal judgments. Judea Pearl (2000) talks about interventions in causal models as "oracles for causal inference", showing that a natural way to characterise many causal inferences is as the outcome of self-interrogation by virtual intervention. This idea has helped make sense of counterfactual reasoning and responsibility attribution (Gerstenberg et al., 2013, 2015). Furthermore, a number of recent papers propose that theory-based inference about the social and physical world are rooted in mental simulation (Hamrick et al., 2016, 2015; Smith & Vul, 2014). Again, the problem of setting up and running such internal simulations seems closely related to the problem of choosing how to experiment with the actual world.

Conceptions of executive control and metacognition mirror conceptions of active learning, with both characterised by perturbing otherwise autonomous processes (Brass & Haggard, 2007), selective attention (Broadbent, 1970; Shiffrin & Schneider, 1977), and planning (Norman & Shallice, 1986). Along these lines, a number of philosophers have suggested an intimate relationship between an interrogatory control loop and consciousness (Baars, 1997; Dennett, 1991; Hofstadter, 1980, 2007). For example, Daniel Dennett proposes that executive control has its evolutionary origins in the gradual internalisation of question asking. Dennett's "just-so" story involves the early Homo sapien asking protolinguistic questions of his peers. Upon asking a question and finding no-one around to answer it, the early-man realises he knows the answer himself. Over time, he learns to ask himself questions surreptitiously, eventually internalising the process as a control loop through which he can interrogate himself and so outwit his peers. The idea that

we gradually internalise our ability to intervene suggests a similar story, but unlike Dennett's, it does not depend on language. On this view, the learner's hand that reaches into the world throughout this thesis might be, in some sense, the same hand that steers *Neurath's ship*, directing attention and resources, testing things out and tinkering with the models in the head. This places active causal learning at the front and centre of cognition as a near cousin, or even a prerequisite for the development of executive control and consciousness.

9.4 Conclusions

In summary, this thesis examined active causal learning over ten experiments and developed process accounts of both complex structure learning and intervention selection. Chapter 3 showed that intervention and causal judgment are well described by probabilistic inference and information-driven testing constrained by forgetting old evidence and conservatism. Chapter 4 examined children's interventional causal structure learning, applying an information theoretic framework and demonstrating a developmental improvement in interventional learning. Chapter 5 developed a process level account of structure change in the face of a generally unmanageably large hypothesis space. The *Neurath's ship* model captured participants' sequences of judgments by assuming that they explored a sequence of local changes to a single global working hypothesis. In parallel, Chapter 5 developed a locally focused intervention scheme, based on the assumption that learners — unable to consider the whole hypothesis space — focused their interventions on resolving manageable sub-problems, so supporting their limited learning trajectories. Chapter 6 focused on the role of time in causal learning, examining people's structural inferences based on repeated observations of the components of a causal device activating over time. This chapter modelled the roles of temporal order and timing information and contains the first demonstration that people can use delay variability alone to distinguish direct and indirect causation. Chapter 7 built on the framework from Chapter 6 to explore interventional learning in the continuous time context. It showed that people organised their interventions in time as well as across components in ways that created approximately independent trials and learning that we could again understand as based in the incremental construction of a global model. Finally, Chapter 8 looked at active learning in dynamic physical microworlds, where interventions involved extended interactions with objects in physical space. We found that active learners came up with creative "experiments" that revealed the worlds' hidden

properties effectively, but that yoked learners did not inherit these advantages. Finally, in Chapter 9, I discussed extensions to the *Neurath's ship* model of causal theory change and active causal learning in cognition more generally, proposing in particular a close relationship between strategic control in the world and in the head. In sum, the story of complex cognition begins and ends the same way, with intervention as the mechanism of discovery and causality the subject matter. It is through intervention that we learn and control the causal world, but also the virtual world we have constructed.

Appendix A

Further formalizing Neurath's ship

Chapter Section 3.1 provided a formal framework for modelling active learning and intervention selection. Here we provide the equations for the more general case of inference without parameter knowledge as used in Chapter 5, Experiment 5: Unknown Strengths.

A.1 Representation and inference

A noisy-OR parametrized causal model m over variables X, with strength and background parameters w_S and w_B assigns a likelihood to each datum (a complete observation, or the outcome of an intervention) $\mathbf{d}$ as the product of the probability of each variable that was not intervened upon given the states of its parents

$$P(\mathbf{d}|m, \mathbf{w}) = \prod_{x \in X} P(x|\mathbf{d}_{pa(x)}, \mathbf{w}) \tag{A.1}$$

$$P(x|\mathbf{d}_{pa(x)}, \mathbf{w}) = x + (1 - 2x)(1 - w_B)(1 - w_S)^{\sum_{y \in pa(x)} y} \tag{A.2}$$

where $pa(x)$ denotes the parents of variable x in the causal model (see Figure 3.1.1 for an example). We can thus compute the posterior probability of model $m \in M$ over a set of models $\mathcal{M}$ given a prior $P(M)$ and data $D = \{\mathbf{d}^i\}$ associated with interventions $C = \{\mathbf{c}^i\}$. We can condition on w_S and w_B if known (e.g. in Experiment 1–4)

$$P(m|D,\mathbf{w}) = \frac{P(D|m,\mathbf{w};C)P(m)}{\sum_{m'\in M} P(D|m',\mathbf{w};C)P(m')} \tag{A.3}$$

or else marginalize over their possible values (e.g. in Experiment 5)

$$P(m|D) = \frac{\int_{\mathbf{w}} P(D|m,\mathbf{w};C)p(\mathbf{w})P(m)\ \mathrm{d}\mathbf{w}}{\sum_{m'\in M}\int_{\mathbf{w}} P(D|m',\mathbf{w};C)p(\mathbf{w})P(m')\ \mathrm{d}\mathbf{w}} \tag{A.4}$$

A.2 Intervention choice

The value of an intervention can be quantified relative to a notion of uncertainty. Here we adopt Shannon entropy (Shannon, 1951), for which the uncertainty in a distribution over causal models M is given by

$$H(M) = -\sum_{m\in M} P(m)\log_2 P(m) \tag{A.5}$$

Assuming $\mathbf{w}$ is known, let $\Delta H(M|\mathbf{d},\mathbf{w};\mathbf{c})$ refer to the reduction in uncertainty going from prior $P(M)$ to posterior $P(M|\mathbf{d},\mathbf{w};\mathbf{c})$ after performing intervention c, then seeing data $\mathbf{d}$

$$\Delta H(M|d,\mathbf{w};\mathbf{c}) = \left[-\sum_{m\in M} P(m)\log P(m)\right] - \left[-\sum_{m\in M} P(m|\mathbf{d},\mathbf{w};\mathbf{c})\log P(m|\mathbf{d},\mathbf{w};\mathbf{c})\right] \tag{A.6}$$

Given this objective, we can define the value of an intervention as the expected reduction in uncertainty after seeing its outcome. To get the expectancy, we must average, prospectively, over the different possible outcomes $\mathbf{d} \in \mathcal{D}_{\mathbf{c}}$ (where $\mathcal{D}_{\mathbf{c}}$ is the space of possible outcomes of intervention $\mathbf{c}$) weighted by their marginal likelihoods under the prior, giving

$$\mathop{\mathbb{E}}_{\mathbf{d}\in\mathcal{D}_{\mathbf{c}}}\left[\Delta H(M|\mathbf{d},\mathbf{w};\mathbf{c})\right] = \sum_{\mathbf{d}\in\mathcal{D}_{\mathbf{c}}}\left(\Delta H(M|\mathbf{d},\mathbf{w};\mathbf{c})\sum_{m\in M} P(\mathbf{d}|m,\mathbf{w};\mathbf{c})P(m)\right) \tag{A.7}$$

For a greedily optimal sequence of interventions $\mathbf{c}^1,\ldots,\mathbf{c}^t$, we take $P(M|D^{t-1},\mathbf{w};C^{t-1})$ as $P(M)$ and $P(M|D^t,\mathbf{w};C^{t-1},\mathbf{c}^t)$ as $P(M|\mathbf{d},\mathbf{w};\mathbf{c})$ in Equation A.6. The most valuable intervention at a given time point is then

$$\mathbf{c}^t = \arg\max_{\mathbf{c}\in\mathcal{C}} \mathop{\mathbb{E}}_{\mathbf{d}\in\mathcal{D}_\mathbf{c}} \left[\Delta H(M|\mathbf{d}, D^{t-1}, \mathbf{w}; C^{t-1}, \mathbf{c})\right] \tag{A.8}$$

If $\mathbf{w}$ is unknown, we must use the marginal distribution, replacing Equation A.6 with

$$\Delta H(M|d;\mathbf{c}) = \left[-\sum_{m\in M} P(m)\log P(m)\right] - \left[-\sum_{m\in M}\int_{\mathbf{w}} P(m|\mathbf{d},\mathbf{w};\mathbf{c})p(\mathbf{w})\ \mathrm{d}\mathbf{w} \log \int_{\mathbf{w}} P(m|\mathbf{d},\mathbf{w};\mathbf{c})p(\mathbf{w})\ \mathrm{d}\mathbf{w}\right] \tag{A.9}$$

A.3 An algorithmic-level model of sequential belief change

Let E be an adjacency matrix such that the upper triangle entries where E_{ij} (if $i < j \leq N$) denotes the state of edge $i - j$ in a causal model m. Any model $m \in M$ corresponds to a setting for all E_{ij} where $i < j \leq N$, to one of three edge states $e \in \{1 : i \to j,\ 0 : i \nrightarrow j,\ -1 : i \leftarrow j\}$. By starting with any hypothesis and iteratively sampling from the conditional distributions on edge states $P(E_{ij}|E_{\setminus ij}, \mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t)$ (Goudie & Mukherjee, 2011) using the following equation:

$$P(E_{ij} = e|E_{\setminus ij}, \mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t) = \frac{P(E_{ij} = e|E_{\setminus ij}, \mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t)}{\sum_{e'\in E_{ij}} P(E_{ij} = e'|E_{\setminus ij}, \mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t)} \tag{A.10}$$

we can cheaply generate chains of dependent samples from $P(M|\mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t)$. This can be done systematically (cycling through all edges $\in i < j \leq N$), or randomly selecting the next edge sample with $P(\frac{1}{|i,j|})$ where $|i, j|$ is the number of edges in the graph. Here we assume random sampling for simplicity. Thus, on each step, the selected E_{ij} is updated using the newest values of $E_{\setminus ij}$.[1] Specifically, we assume that after each new piece of evidence arrives:

1. The learner begins sampling with edges $E_{ij}^{(0)}$ for all i and j set as they were in their previous judgment b^{t-1}.
2. They then randomly select an edge E_{ij} in $i < j \leq N$ to update.
3. They resample $E_{ij}^{(1)}$ using Equation 5.1.

[1] Edge changes that would create a cyclic graph always have a probability of zero

4. If the search does not result in a new model they keep collecting evidence $\mathcal{D}_r^t = \{\mathcal{D}_r^t, \mathbf{d}^t\}$, $\mathbf{c} = \{\mathcal{C}_r^t, \mathbf{c}^t\}$. If it does, the evidence is used up and forgotten and they begin collecting evidence again (e.g. resetting $\mathcal{D}_r^t = \{\}$ and $\mathcal{C}_r^t = \{\}$).

5. The learner repeats steps 1 to 4 k times, with their final edge choices $E^{(k)}$ constituting their new belief b^t.

We assume for simplicity that b^0, before any data has been seen is an unconnected graph, but have tested this assumption by fitting the data from t=2 onward only finding better fits overall and a stronger win for *Neurath's ship* over the other models we consider.

Resampling, hill climbing or random change

We also consider generalizations of Equation A.10 allowing transitions to be governed by higher powers of $P(E_{ij} = e|E_{\setminus ij}, \mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t)$

$$P^\omega(E_{ij} = e|E_{\setminus ij}, \mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t) = \frac{P^\omega(E_{ij} = e|E_{\setminus ij}, \mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t)}{\sum_{e' \in E_{ij}} P^\omega(E_{ij} = e'|E_{\setminus ij}, \mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t)} \tag{A.11}$$

yielding stronger preference for the most likely state of e_{ij} if $\omega > 1$ and more random sampling if $\omega < 1$.

A distribution over search lengths

We assume that for each update, the learner's length of search k is drawn from a Poisson distribution with average $\lambda \in [0, \infty]$

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{A.12}$$

Putting these together

To calculate the probability distribution of new belief b^t given $\mathbf{d}^t$, b^{t-1} search behavior ω and a chain of length k, we first construct the transition matrix R_t^ω for the Markov search chain by averaging over the conditional distributions associated with the choice of each edge, weighted by the probability of selecting that edge

$$R_t^{\omega} = \sum_{i<j\leq N} P^{\omega}(E_{ij} = e|E_{\setminus ij}, \mathcal{D}_r^t, \mathbf{w}; \mathcal{C}_r^t) \times \frac{1}{|i,j|} \tag{A.13}$$

for each possible belief b.

By raising this transition matrix to the power k (i.e. some search length) and selecting the row corresponding to starting belief $[(R_t^{\omega})^k]_{b^{t-1}}$, we get the probability of adopting each $m \in M$ as new belief b^t (see Figure 5.1 for a visualization) at the end of the k length search

$$P(B^t|\mathcal{D}_r^t, b^{t-1}, \omega, k; \mathcal{C}_r^t) = [(R_t^{\omega})^k]_{b^{t-1}m} \tag{A.14}$$

Finally, by averaging over different possible chain lengths k, weighted by their probability Poisson(λ) we get the marginal probability that a learner will move to each possible new belief in B at t

$$P(B^t|\mathcal{D}_r^t, b^{t-1}, \omega, \lambda; \mathcal{C}_r^t) = \sum_0^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} [(R_t^{\omega})^k]_{b^{t-1}m} \tag{A.15}$$

A.4 A local uncertainty schema

Edge focus

Relative to a focus on an edge E_{xy}, intervention values were calculated using expected information as in Appendix A, but assuming prior entropy as that of a uniform distribution over the three possible edge states

$$H(E_{xy}|E_{\setminus xy}) = -3\left(\frac{1}{3}\log_2\frac{1}{3}\right) \tag{A.16}$$

and calculating posterior entropies for the possible outcomes $\mathbf{d} \in \mathcal{D}$ using

$$H(E_{xy}|E_{\setminus xy}, \mathbf{d}, \mathbf{w}; \mathbf{c}) = -\sum_{z\in\{-1,0,1\}} P(E_{xy} = z|E_{\setminus xy}, \mathbf{d}, \mathbf{w}; \mathbf{c}) \log_2 P(E_{xy} = z|E_{\setminus xy}, \mathbf{d}, \mathbf{w}; \mathbf{c}) \tag{A.17}$$

Effect focus entropy

Relative to a focus on the effects of variable x, intervention values were calculated using expected information as in Appendix A but using prior entropy, calculated by partitioning a uniform prior over models M into sets of models $\mathrm{Mo}(z)$ corresponding to each descendant set $z \subseteq \mathrm{De}(x)$

$$H(\mathrm{De}(x)) = -\sum_{z \subseteq \mathrm{De}(x)} \left(\sum_{m \in \mathrm{Mo}(z)} \frac{1}{|M|} \right) \log_2 \left(\sum_{m \in \mathrm{Mo}(z)} \frac{1}{|M|} \right) \tag{A.18}$$

Posterior entropies were then calculated by summing over probabilities of the elements in each $\mathrm{Mo}(z)$ for each $z \subseteq \mathrm{De}(x)$

$$H(\mathrm{De}(x)|\mathbf{d}, \mathbf{w}; \mathbf{c}) = -\sum_{z \subseteq \mathrm{De}(x)} \left(\sum_{m \in \mathrm{Mo}(z)} P(m|\mathbf{d}, \mathbf{w}; \mathbf{c}) \right) \log_2 \left(\sum_{m \in \mathrm{Mo}(z)} P(m|\mathbf{d}, \mathbf{w}; \mathbf{c}) \right) \tag{A.19}$$

Confirmation focus entropy

Relative to a focus on distinguishing current hypothesis b^t from null hypothesis b^0, intervention values were calculated using expected information as above but prior entropy was always based on a uniform prior over the two hypotheses

$$H(\{b^t, b^0\}) = -2 \left(\frac{1}{2} \log_2 \frac{1}{2} \right) \tag{A.20}$$

and posterior entropies were calculated using

$$H(\{b^t, b^0\}|\mathbf{d}, \mathbf{w}; \mathbf{c}) = -\sum_{z \in \{0,t\}} \frac{P(b^z|\mathbf{d}, \mathbf{w}; \mathbf{c})}{\sum_{z' \in \{0,t\}} P(b^{z'}|\mathbf{d}, \mathbf{w}; \mathbf{c})} \log_2 \frac{P(b^z|\mathbf{d}, \mathbf{w}; \mathbf{c})}{\sum_{z' \in \{0,t\}} P(b^{z'}|\mathbf{d}, \mathbf{w}; \mathbf{c})} \tag{A.21}$$

A.5 Additional modelling details

All models were fit using maximum likelihood. Maximum likelihood estimates were found using Brent (for one parameter) or Nelder-Mead (for several parameters) optimization, as implemented by R's `optim` function. Convergence to global optima was checked by repeating all optimizations with a range of randomly selected starting parameters.

k For averaging across different values of k in the belief models, we capped k at 50 and renormalized the distribution such that $P(k \geq 0 \wedge k \leq 50) = 1$. This made negligible difference to the fits since the probabilities of $P(B^t|\mathbf{d}^r, b^{t-1}, \omega, k; \mathcal{C}_r^t)$ for values of $k \gg N$ (where N is the number of variables) were very similar.

ϵ To allow that participants are liable to occasionally lapse concentration or forget the outcome of a test, we included a lapse parameter ϵ – i.e., a parametric amount of decision noise $\epsilon \in [0, 1]$ – so that the probability of a belief would be a mixture of that predicted by the model and uniform noise. This ensured that occasional random judgments did not have undue effects on the other parameters of each model.

b^0 We assume for simplicity that people's starting belief, b^0, before any data has been seen, is an unconnected graph.

Marginalization

For all modelling in Experiment 3, we had to average over the unknown noise $\mathbf{w}$. To do this, we drew 1000 paired uniformly distributed w_S and w_B samples and averaged over these when computing marginal likelihoods and posteriors. These marginal priors and posteriors were used for computing expected information gain values.

Evaluating fits

Baseline acts as the null model for computing BIC's (Schwarz, 1978) and pseudo-R^2's (Dobson, 2010) for all other models.

A.6 Model recovery

TABLE A.1: Belief Model Recovery Analysis

Bramley, Lagnado, and Speekenbrink (2015)

All simulated participants

	Baseline	Rational	WSLS	SE	NS-RE	NS
Baseline	**134**	0	4	0	1	0
Rational	39	**95**	4	0	0	1
WSLS	5	0	**133**	1	0	0
SE	1	0	2	**119**	17	0
NS-RE	2	0	1	53	**82**	1
NS	3	0	3	11	**72**	50

Best fit participants

	Baseline	Rational	WSLS	SE	NS-RE	NS
Random	**1**	0	0	0	0	0
Rational	0	**10**	0	0	0	0
WSLS	0	0	**56**	0	0	0
SE	0	0	0	**13**	2	0
NS-RE	0	0	1	9	**20**	0
NS	0	0	1	1	**16**	9

Exp 1: Learning larger models

All simulated participants

	Baseline	Rational	WSLS	SE	NS-RE	NS
Baseline	**114**	1	3	0	2	0
Rational	42	**75**	2	0	1	0
WSLS	4	1	**115**	0	0	0
SE	6	0	9	**94**	11	0
NS-RE	2	0	1	3	**112**	2
NS	3	0	3	2	30	**82**

Best fit participants

	Baseline	Rational	WSLS	SE	NS-RE	NS
Baseline	**1**	0	0	0	0	0
Rational	0	0	0	0	0	0
WSLS	1	1	**27**	0	0	0
SE	0	0	1	**31**	0	0
NS-RE	0	0	0	1	**29**	1
NS	0	0	0	0	4	**23**

Exp 2: Unknown strengths

All simulated participants

	Baseline	Rational	WSLS	SE	NS-RE	NS
Baseline	**107**	1	3	0	0	0
Rational	17	**92**	2	0	0	0
WSLS	7	0	**100**	2	2	0
SE	9	0	6	**93**	3	0
NS-RE	9	0	8	3	**89**	2
NS	5	2	9	2	**70**	23

Best fit participants

	Baseline	Rational	WSLS	SE	NS-RE	NS
Baseline	**2**	0	0	0	0	0
Rational	1	**18**	0	0	0	0
WSLS	1	0	**16**	0	1	0
SE	0	0	0	**30**	0	0
NS-RE	0	0	0	0	**5**	0
NS	0	2	3	0	**22**	10

Note: Rows denote simulation rule and columns the model used to fit the simulated choices. The number in each cell shows how many of the simulations using this rule were best fit by that model. Right hand side restricts this to simulations using the parameters taken from participants who were actually best fit by each model.

Appendix B

Representing time

B.1 Collider likelihood

Pooled model

For the Collider in Experiments 6 and 7, event E happens as the two causal influences of A and B arrived (i.e., conjunctive common-effect; see Equation 6.4). Thus, the observed between-event intervals t_{AE} and t_{BE} may contain waiting time and so do not necessarily reflect the underlying causal delays $t_{A\to E}$ and $t_{B\to E}$ as we have assumed for the other structures. To model the joint likelihood of the two observed intervals, we have to discriminate two cases: Either (1) the causal influence of B was waiting for the influence of A and therefore E happened as the delay of A arrived (i.e., $t_{AE} = t_{A\to E}$ but $t_{BE} \geq t_{B\to E}$) or (2) the causal influence of A was waiting for the influence of B to arrive and E happened as the delay of B arrived (i.e., $t_{BE} = t_{B\to E}$ but $t_{AE} \geq t_{A\to E}$).

Let the influence of B waiting for A (i.e., Case 1). In this case, the joint likelihood is given by the gamma likelihood of t_{AE} (as t_{AE} does in fact equal $t_{A\to E}$ and is therefore gamma distributed) weighted by the probability of t_{BE} being in fact larger than the respective gamma distributed event $t_{B\to E}$. As we assume the same parameters α and μ for both links (pooled model), the likelihood can be written as

$$p(t_{AE}, t_{BE}|\alpha, \mu; t_{AE} = t_{A\to E}, t_{BE} \geq t_{B\to E}) = p(t_{AE}|\alpha, \mu) \cdot p(t_{BE} \geq t_{B\to E}|\alpha, \mu) \quad \text{(B.1)}$$

Analogously, for the case in which A is waiting for B (i.e., Case 2) it holds

$$p(t_{AE}, t_{BE}|\alpha, \mu; t_{AE} \geq t_{A\to E}, t_{BE} = t_{B\to E}) = p(t_{BE}|\alpha, \mu) \cdot p(t_{AE} \geq t_{A\to E}|\alpha, \mu) \quad \text{(B.2)}$$

As both cases are mutual exclusive and therefore constitute a partitioning of the joint likelihood, the joint likelihood can be written as a sum of both (law of total probability)

$$p(t_{AE}, t_{BE}|\alpha, \mu) = p(t_{AE}, t_{BE}|\alpha, \mu; t_{AE} = t_{A\to E}, t_{BE} \geq t_{B\to E}) + p(t_{AE}, t_{BE}|\alpha, \mu; t_{AE} \geq t_{A\to E}, t_{BE} = t_{B\to E}) \quad \text{(B.3)}$$

$$= p(t_{AE}|\alpha, \mu) \cdot p(t_{BE} \geq t_{B\to E}|\alpha, \mu) + p(t_{BE}|\alpha, \mu) \cdot p(t_{AE} \geq t_{A\to E}|\alpha, \mu) \quad \text{(B.4)}$$

with $p(t_{AE}|\alpha, \mu)$ and $p(t_{BE}|\alpha, \mu)$ being gamma distributed and $p(t_{AE} \geq t_{A\to E}|\alpha, \mu)$ and $p(t_{BE} \geq t_{B\to E}|\alpha, \mu)$ following the gamma's cumulative distribution function with

$$p(t_{AE} \geq t_{A\to E}|\alpha, \mu) = \int_0^{t_{AE}} \frac{(\frac{\alpha}{\mu})^\alpha}{\Gamma(\alpha)} (x)^{\alpha-1} e^{-\frac{\alpha}{\mu}x}\, dx \quad \text{(B.5)}$$

and for $p(t_{BE} \geq t_{B\to E}|\alpha, \mu)$ analogously.

Independent model

In the independent model, each causal connection between a variable X and its effect Y is assumed to have its own set of parameters α_{XY} and μ_{XY}. Therefore, the Collider likelihood in the independent model is given by

$$\begin{aligned} p(t_{AE}, t_{BE}|\alpha_{AE}, \alpha_{BE}, \mu_{AE}, \mu_{BE}) =& p(t_{AE}|\alpha_{AE}, \mu_{AE}) \cdot p(t_{BE} \geq t_{B\to E}|\alpha_{BE}, \mu_{BE}) \\ &+ p(t_{BE}|\alpha_{BE}, \mu_{BE}) \cdot p(t_{AE} \geq t_{A\to E}|\alpha_{AE}, \mu_{AE}) \end{aligned} \quad \text{(B.6)}$$

Disjunctive Collider

In our experiments, we used conjunctive Colliders. However, in other scenarios a disjunctive combination function of the causal influences may be more natural. In this case, the activation time of effect event E is determined by the first arrival of the causes' influences

$$t_E = \min[t_A + t_{A\to E}, t_B + t_{B\to E}] \tag{B.7}$$

In this case, one of the underlying causal delays $t_{A\to E}$ or $t_{B\to E}$ is overshadowed by E's happening resulting in a smaller observed delay. Analogously to the conjunctive Collider, there are two cases: (1) the influence of A arrives first, causing E to happen and overshadowing the influence of B (i.e., $t_{AE} = t_{A\to E}$ but $t_{BE} \leq t_{B\to E}$) and (2) the influence of B arrives first overshadowing the influence of A (i.e., $t_{BE} = t_{B\to E}$ but $t_{AE} \leq t_{A\to E}$). Thus, the joint likelihood of a disjunctive (pooled delay) Collider can be written as

$$p(t_{AE}, t_{BE}|\alpha, \mu) = p(t_{AE}|\alpha, \mu) \cdot p(t_{BE} \leq t_{B\to E}|\alpha, \mu) + p(t_{BE}|\alpha, \mu) \cdot p(t_{AE} \leq t_{A\to E}|\alpha, \mu) \tag{B.8}$$

$$= p(t_{AE}|\alpha, \mu) \cdot (1 - p(t_{BE} \geq t_{B\to E}|\alpha, \mu)) + p(t_{BE}|\alpha, \mu) \cdot (1 - p(t_{AE} \geq t_{A\to E}|\alpha, \mu)) \tag{B.9}$$

B.2 Simple Monte Carlo estimation: Experiment 6 and 7

As there is no closed form solution for the marginal likelihoods $p(\mathbf{d}|m)$ of data $\mathbf{d}$ under structure m, we used a simple Monte Carlo sampling scheme to approximate the multiple integral. For this purpose, we drew $B = 100,000$ independent samples from the respective parameters' prior distributions $p(\lambda|m)$, $p(\alpha|m)$ and $p(\mu|m)$ and averaged over the likelihoods (see Equation 6.9) at the sampled points in parameter space

$$p(\mathbf{d}|m) = \int p(\mathbf{d}|\lambda, \alpha, \mu; m) \cdot p(\lambda|m) \cdot p(\alpha|m) \cdot p(\mu|m)\ d\lambda\ d\alpha\ d\mu \tag{B.10}$$

$$= \frac{1}{B} \sum_{b=1}^{B} p(\mathbf{d}|\lambda^{(b)}, \alpha^{(b)}, \mu^{(b)}; m) \tag{B.11}$$

with $\lambda^{(b)}$, $\alpha^{(b)}$, and $\mu^{(b)}$ being the b's sampled points from the prior distributions.

B.3 Markov Chain Monte Carlo estimation: Experiment 8

In Experiment 8, we could use an uninformative prior for the parameters of the gamma distribution (as no collider was involved). For one causal link and the gamma's (α, θ) parametrization with $\mu = \frac{\alpha}{\theta}$, we can derive the posterior based on a conjugate prior assuming "no prior observations"

$$p(\alpha, \theta|\mathbf{d}; m) \propto \frac{p^{\alpha-1}e^{-\frac{q}{\theta}}}{\Gamma(\alpha)^n \theta^{\alpha n}} \tag{B.12}$$

for n data points $\mathbf{d}$ with $p = \prod d_i$ and $q = \sum d_i$.[1] The normalizing constant of the equation's right hand side is our target of interest, namely the marginal likelihood of the data given the structure of interest $p(\mathbf{d}|m)$. To approximate the integral, we used a two-step procedure:

1. We generated a sample from the posterior over α and θ via the Metropolis–Hastings algorithm (i.e., MCMC) with 10,000 points sampled from 10 chains each with Gaussian proposal distribution on α ($SD = 10$) and θ ($SD = 5$) and burn-in of 1,000 and only each tenth point taken (i.e., thinning). We run the sampler ten times to check for convergence (see Gelman, Carlin, Stern, & Rubin, 2004).

2. We used the obtained sample to estimate the marginal likelihood with the method proposed by Chib and Jeliazkov (2001). Although the method formally works with just one sampled point, we used a subset to generate a more stable estimate.We randomly drew 1,000 points from the MCMC sample and took the 50 points with the largest likelihoods in this subsample. For each of these points, we calculated the marginal likelihood estimate with the method proposed by Chib and Jeliazkov (2001) and averaged over these to get our estimate of $p(\mathbf{d}|m)$.

B.4 Checking sensitivity to priors

We can assess the sensitivity of the Simple Monte Carlo model predictions to prior choices by comparing them to the predictions of the Markov Chain Monte Carlo procedure

[1] Note that we describe delays in terms of their shape α and mean μ in the main text to aid exposition. However, in statistical applications including approximating inference it is more common and more convenient to work with shape and rate θ.

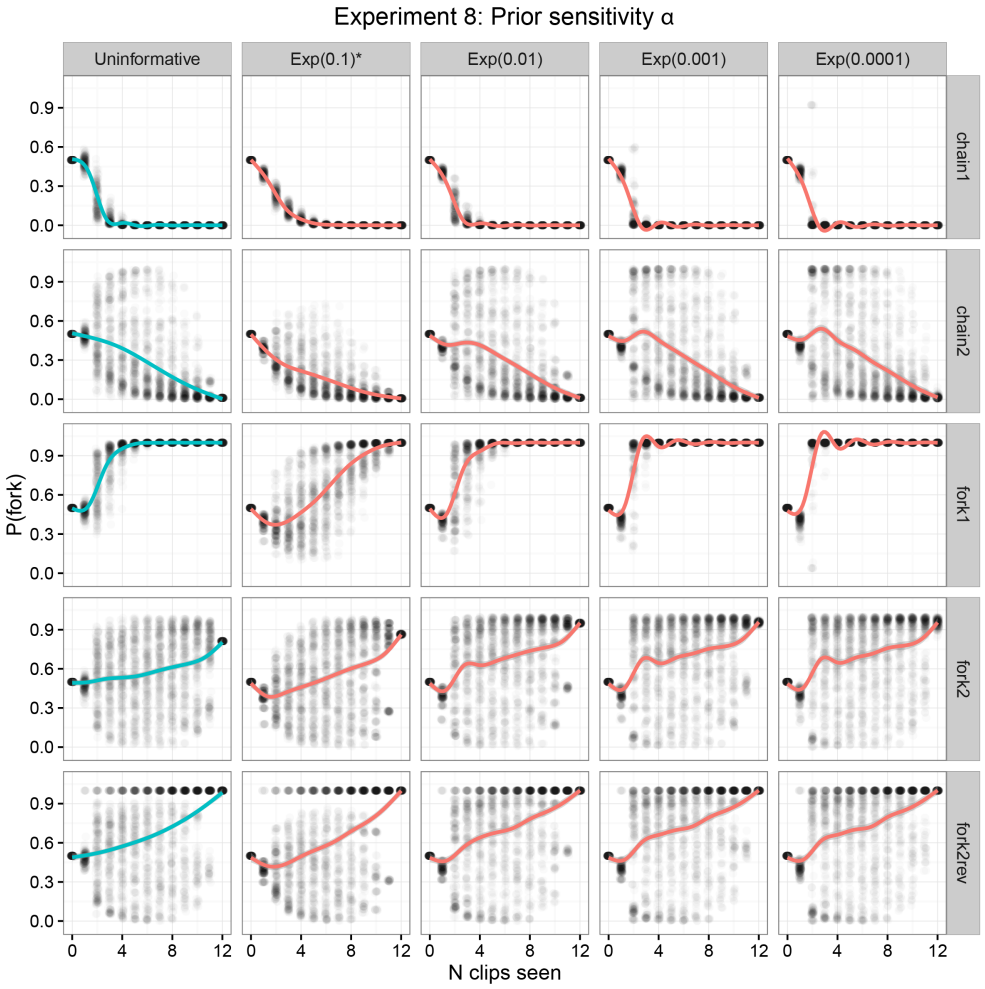


FIGURE B.1: Sensitivity of α prior on model predictions in Experiment 8. Left hand column (teal line) shows predictions using an "improper" uninformative prior. Other columns show predictions under different priors on α. Asterisk indicates the values used for Experiment 6 and 7. The prior on μ for these simulations was Exponential(0.0001). As in Figure 6.13, individual points are for subsets of the tests seen by different participants at different points during the experiment.

we used to estimate posteriors in Experiment 8. The Markov Chain procedure gives posterior predictions based on an uninformative "improper" (Hartigan, 2012) prior but cannot be used for the Collider structure in Experiment 6 and 7. We see in Figures B.1 and B.2 that there is a little sensitivity to choice of priors on α and μ. Particularly, too high a rate for μ leads to an initial preference for shorter delays and hence the chain under which the delays are necessarily shorter. Additionally, too low a rate for either α or μ led to less stable predictions as few samples fall in the range of the true generative model. However, our chosen values of 0.1 for α and 0.0001 for μ make these effects negligible for the range of event timings we consider in Experiments 6–8.

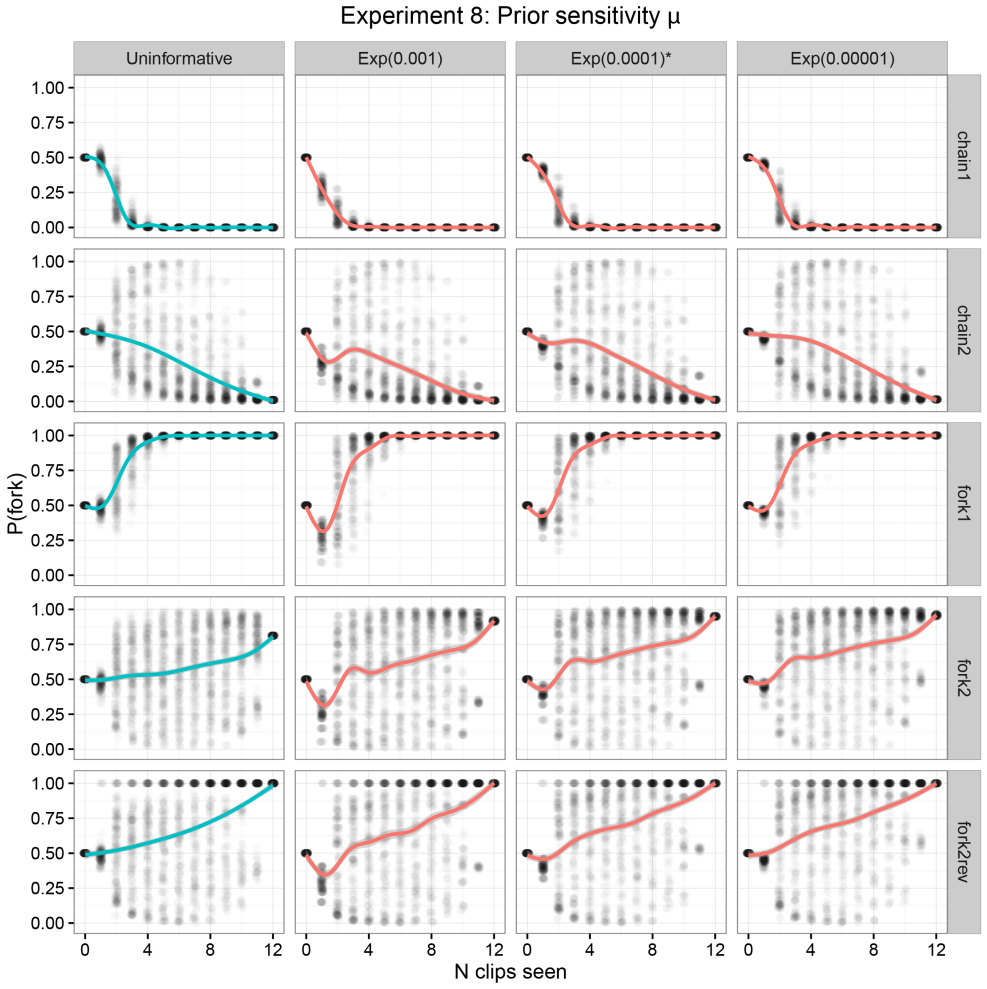


FIGURE B.2: Sensitivity of μ prior on model predictions in Experiment 8. Left hand column (teal line) shows predictions using an "improper" uninformative prior. Other columns show predictions under different priors on μ. Asterisk indicates the values used for Experiment 6 and 7. The prior on α for these simulations was Exponential(0.1).

B.5 Correlations Among Heuristic Predictors in Experiment 8

TABLE B.1: Experiment 8: Correlation Between Heuristic Measures

First Judgment (after 6 tests)							
	$\text{cor}(t_{SA}, t_{SB})$	$\sigma(t_{SA})$	$\sigma(t_{SB})$	$\sigma(t_{AB})$	$\text{APD}(t_{SA})$	$\text{APD}(t_{SB})$	$\text{APD}(t_{AB})$
Delay_I	0.69	-0.03	0.42	-0.31	-0.01	0.52	-0.44
$\text{cor}(t_{SA}, t_{SB})$		-0.02	0.13	-0.66	0.01	0.23	-0.68
$\sigma(t_{SA})$			0.15	0.16	0.69	0.12	0.13
$\sigma(t_{SB})$				0.51	0.16	0.90	0.42
$\sigma(t_{AB})$					0.12	0.39	0.88
$\text{APD}(t_{SA})$						0.20	0.17
$\text{APD}(t_{SB})$							0.33
Second and Third Judgments (after all 12 tests)							
	$\text{cor}(t_{SA}, t_{SB})$	$\sigma(t_{SA})$	$\sigma(t_{SB})$	$\sigma(t_{AB})$	$\text{APD}(t_{SA})$	$\text{APD}(t_{SB})$	$\text{APD}(t_{AB})$
Delay_I	0.97	–	0.38	-0.61	0.04	0.45	-0.63
$\text{cor}(t_{SA}, t_{SB})$			0.24	-0.70	0.04	0.37	-0.77
$\sigma(t_{SA})$			–	–	–	–	–
$\sigma(t_{SB})$				0.51	0.03	0.93	0.38
$\sigma(t_{AB})$					-0.01	0.37	0.92
$\text{APD}(t_{SA})$						0.12	0.05
$\text{APD}(t_{SB})$							0.24

Note: In Experiment 8, $\sigma(t_{SA})$ was the same for all devices so it is uncorrelated with all measures once all evidence has been observed.

B.6 Estimating posteriors in Experiment 9

In order to calculate normative Bayesian predictions for model judgments in Experiment 9, based on knowledge of the parameters α, μ, w_S, data $\mathbf{d}$ and interventions $\mathbf{c}$, we need to calculate the likelihood for every possible pattern of actual causation under each model $m \in \mathcal{M}$ (we call these $\mathbf{z} \in \mathbf{Z}_{\mathbf{d}}^m$) summing over them to get a marginal likelihood for $p(\mathbf{d}|m, \mathbf{c})$. We can calculate the likelihood of a specific path $\mathbf{z}$ as the product of the likelihood of each event occurring at the time it did, given when its parent occurred in $\mathbf{z}$, combined with the failure-likelihood of any non-occurring events that the model predicted should have occurred but did not occur in $\mathbf{z}$ $\phi \in \Phi_{\mathbf{z}}$. These failures could either be due to the $1 - w_S$ causal failure rate, or due to the effect simply failing to occur before the end of the $[0, \tau]$ observational window. Similarly to Equation 6.3 in Chapter 6, we can compute the likelihood of each particular event occurring at time $d_X^{(i)}$ given its parent in $\mathbf{z}$ — $\text{pa}_{\mathbf{z}}(d_X^{(i)})$ — occurred at time $t_{\mathbf{z}}^{\text{pa}(i)}$ and parameters $\mathbf{w} = [\alpha, \mu, w_S]$. This is the product of the likelihood of delay $d_X^{(i)} - t_{\mathbf{z}}^{\text{pa}(i)}$ multiplied by the w_S probability that the

causal connection worked, giving

$$p(d_X^{(i)}|\operatorname*{pa}_{\mathbf{z}}(d_X^{(i)}),\tau,\mathbf{w}) = w_S \frac{(\frac{\alpha}{\mu})^\alpha}{\Gamma(\alpha)} \left(d_X^{(i)} - t_{\mathbf{z}}^{\mathrm{pa}(i)}\right)^{\alpha-1} e^{-\frac{\alpha}{\mu}\left(d_X^{(i)} - t_{\mathbf{z}}^{\mathrm{pa}(i)}\right)}. \tag{B.13}$$

We must also consider the probability of any unexpected non-events $\phi \in \Phi_{\mathbf{z}}$. How surprising these are depends on how close to the end of the observational period the would-have-been parent $\mathrm{pa}_{\mathbf{z}}(\phi)$ occurred (e.g. $t_{\mathbf{z}}^{\mathrm{pa}(\phi)}$). We can compute these probabilities by adding the chance that the event does not occur at all, to the probability that ϕ is still to occur after the end of the observational window. We can write this as

$$p(\phi|\operatorname*{pa}_{\mathbf{z}}(\phi),\tau,\mathbf{w}) = (1-w_S) + p(\phi|\operatorname*{pa}_{\mathbf{z}}(\phi),\tau,\mathbf{w})w_S \tag{B.14}$$

where $p(\phi > \tau|\,\mathrm{pa}_{\mathbf{z}}(\phi),\mathbf{w})$ is the complement of the cumulative distribution for the Gamma delay function offset by the timing of the parent:

$$p(\phi > \tau|\operatorname*{pa}_{\mathbf{z}}(\phi),\tau,\mathbf{w}) = \frac{1}{\Gamma(\alpha)}\Gamma(\alpha, \frac{\alpha}{\mu}(\tau - t_{\mathbf{z}}^{\mathrm{pa}(\phi)})). \tag{B.15}$$

(e.g. the chance that the activation is still to occur). The likelihood of a particular model and path of actual causation $P(m,\mathbf{z})$ is thus the product of the likelihood of all events that did occur (all $d \in \mathbf{d}$) given their parent in $\mathbf{z}$ (which could be another event in $\mathbf{d}$ or an intervention in $\mathbf{c}$), and the likelihood of the failure of each event that did not occur

$$P(\mathbf{d}|m,\tau,\mathbf{z},\mathbf{w};\mathbf{c}) = \prod_{d^{(i)}\in\mathbf{d}} p(d^{(i)}|\operatorname*{pa}_{\mathbf{z}}(d^{(i)}),\mathbf{w}) \prod_{\phi\in\Phi} p(\phi|\operatorname*{pa}_{\mathbf{z}}(\phi),\mathbf{w}) \tag{B.16}$$

The marginal likelihood of model m is then

$$p(\mathbf{d}|m,\tau,\mathbf{w};\mathbf{c}) = \sum_{\mathbf{z}'\in Z_m^{\mathbf{d}}} p(\mathbf{d},\mathbf{z}'|\tau,\mathbf{w};\mathbf{c}) \tag{B.17}$$

As in the other chapters, a posterior over models $P(M|\mathbf{d},\tau,\mathbf{w};\mathbf{c})$ can then be calculated using Equation 3.3.

Approximating the sum over Z

Our method for calculating the likelihood of the data $\mathbf{d}$ in Experiment 9 required summing over all the possible causal paths $\mathbf{z} \in \mathbf{Z}_{\mathbf{d}}^m$ (e.g. Equation B.17). However, the

cardinality of $\mathbf{Z}_{\mathbf{d}}^{m}$ grows on the order of 2^n. Thus for some trials — especially those for four variable cyclic devices in which there could be hundreds of activations in total — $|\,\mathbf{Z}_{\mathbf{d}}^{m}\,|$ was far too large to consider explicitly. However, the large majority of these paths were extremely unlikely to occur. Thus we used up to three approximations to reduce these to a managable number for each trial. First we always excluded all causal paths that implied delays that were highly implausible (and outwith the range of any delays actually produced by true causal structures in generating the actual evidence). For the *reliable* condition we ruled out causal paths with delays less than 1000ms or more than 2000ms. For the *variable-within* and *variable-between* conditions we ruled out causal paths with delays of less than 150ms or more than 5600ms.

If this still resulted in too many paths to reasonably evaluate, we reduced the number of paths greedily, by removing the least likely cause of the most ambiguous event as a candidate. For example, one event might have 10 candidate causes, while the next most ambiguous have 8. Thus we would remove the least likely of these 10 causes from the pool of candidates, and repeat until the product of all sets of candidate causes per event fell below 2,000.

Finally, on a small number of cyclic device trials where participants intervened rapidly leading to many concurant activation streams, we randomly subsampled $\mathbf{Z}_{\text{sub}}^{m}$ and renormalised our likelihood estimate

$$p(\mathbf{d}|m,\tau;\mathbf{c}) \approx \frac{|\,\mathbf{Z}_{\mathbf{d}}^{m}\,|}{|\,\mathbf{Z}_{\text{sub}}^{m}\,|} \sum_{\mathbf{z}'\in\mathbf{Z}_{m}^{\text{sub}}} p(\mathbf{d},\mathbf{z}'|m,\tau;\mathbf{c}) \tag{B.18}$$

Appendix C

Physics worlds

C.1 Box2D settings

The open source physics engine we used for Experiment 10 is called Box2D, and is available here `https://www.github.com/erincatto/Box2D` with the javascript port available here `https://box2d-js.sourceforge.net`. Our demo code is available here and here: here. After landing at these locations, right click to view the source code.

Below is a list of the Box2D variables and functions as they were defined for Experiment 10: Intervening in space and time:

- Number of steps (frames): 2700
- Trial length = 45s
- Box2D step size: 1/60s ($\approx 17ms$)
- Ratio (pixels to meters): 100 (200 on retina screens)
- Object velocity cap: $30\,{}^{\mathrm{m}}/_{\mathrm{s}}$
- Criterion for refreshing puck locations and velocities: Fastest object is moving at less than $0.25\,{}^{\mathrm{m}}/_{\mathrm{s}}$
- Pause time if locations refreshed = 500ms
- World width: 6m (600 pixels / 1200 on retina screens)
- World height: 4m (400 pixels / 800 on retina screens)
- Global forces: None
- Attractive forces: $+3\,{}^{\mathrm{m}}/_{\mathrm{s}^2}$
- Repulsive forces: $-3\,{}^{\mathrm{m}}/_{\mathrm{s}^2}$
- Controlled object attraction to cursor: $.2 \times \mathrm{distance}(\mathrm{cursor}, \mathrm{controlledobject})\,{}^{\mathrm{m}}/_{\mathrm{s}^2}$

- Controlled object damping: 10[1]
- Puck masses: 1kg (or 2kg for heavy target ball)
- Puck friction: .05[2]
- Puck elasticity: .98
- Puck damping: .05
- Puck radius: .25 m
- Puck object types: Dynamic
- Wall mass: n/a
- Wall friction: .05
- Wall elasticity: .98
- Wall damping: n/a
- Wall width = .2m
- Wall object types: Static

C.2 Online information measures

We defined PD_{mass} as the average predicted divergence between worlds $m \in \mathcal{M}$ differing on the target mass dimension. To write this we split $\mathcal{M}$ into three subsets $\mathcal{M} = \mathcal{M}_A \cup \mathcal{M}_B \cup \mathcal{M}_{\text{same}}$ such that models $\mathcal{M}_A\{i\}, \mathcal{M}_B\{i\}$ and $\mathcal{M}_{\text{same}}\{i\}$ are identical on all dimensions except the target mass. We then evaluated their expected divergence by averaging over all comparisons (e.g. A vs. B, A vs. *same* and B vs. *same*) and all other properties (e.g. $i \in \frac{|\mathcal{M}|}{3}$), using the same Gaussian error assumption as Ullman et al. (2014) and Equation 8.1. To get a measure that increases for greater average divergences (unlike the likelihoods that decreased), we subtracted these scores from 1. The resulting average divergence can be written as:

$$\underset{\text{mass}}{\text{PD}} = 1 - \mathop{\mathbb{E}}_{\mathcal{I}<\mathcal{J}\in[\mathcal{M}_A,\mathcal{M}_B,\mathcal{M}_{\text{same}}]} \left[\mathop{\mathbb{E}}_{i\in\mathcal{I},j\in\mathcal{J}} \left[e^{-\epsilon(\mathbf{x}_i-\mathbf{x}_j)^2} \right] \right]. \tag{C.1}$$

We do the same for PD_{force}, replacing $\mathcal{M}_A, \mathcal{M}_B$ and $\mathcal{M}_{\text{same}}$ with $\mathcal{M}_{\text{attract}}, \mathcal{M}_{\text{repel}}$ and $\mathcal{M}_{\text{none}}$:

[1] Damping in Box2D slows objects while they are not in contact with any other objects (like wind resistance). The controlled object was given high damping so it would not oscillate for a long time around the cursor location.

[2] Friction in Box2D occurs when two objects slide past each other while touching (e.g. a puck sliding along a boundary wall).

$$\underset{\text{force}}{\text{PD}} = 1 - \mathop{\mathbb{E}}_{\mathcal{I}<\mathcal{J}\in[\mathcal{M}_{\text{attract}},\mathcal{M}_{\text{repel}},\mathcal{M}_{\text{none}}]} \left[\mathop{\mathbb{E}}_{i\in\mathcal{I},j\in\mathcal{J}} \left[e^{-\epsilon(\mathbf{x}_i-\mathbf{x}_j)^2} \right] \right]. \tag{C.2}$$

Finally, to compute PD_{any} we repeat this procedure for all 7 dimensions of the problem $\forall$ z $\in Z$ (e.g. the target mass, target force and the five possible distractor forces) and take the average of all of these:

$$\underset{\text{any}}{\text{PD}} = 1 - \mathop{\mathbb{E}}_{z\in Z} \left[\mathop{\mathbb{E}}_{\mathcal{I}<\mathcal{J}\in[\mathcal{M}_{z_1},\mathcal{M}_{z_2},\mathcal{M}_{z_3}]} \left[\mathop{\mathbb{E}}_{i\in\mathcal{I},j\in\mathcal{J}} \left[e^{-\epsilon(\mathbf{x}_i-\mathbf{x}_j)^2} \right] \right] \right]. \tag{C.3}$$

We assumed scaling parameter $\epsilon = 10$ for computing model posteriors (Equation 8.1) and $\epsilon = 10,000$ for the predictive divergence measures (Equations C.1, C.2 and C.3). Using the same parameter for both led to underflow issues since the overall likelihood is based on the product over all 2700 frames, while the predicted divergences are computed separately on every frame. In other respects, trying a range of values for ϵ did not affect the reported comparisons.

References

Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2012). Human memory search as a random walk in a semantic network. In *Advances in Neural Information Processing Systems* (pp. 3050–3058).

Abbott, J. T., & Griffiths, T. L. (2011). Exploring the influence of particle filter parameters on order effects in causal learning. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Agrawal, P., Nair, A., Abbeel, P., Malik, J., & Levine, S. (2016). Learning to poke by poking: Experiential learning of intuitive physics. *arXiv preprint arXiv:1606.07419*.

Ahn, W.-k., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, *54*(3), 299–352.

Aldous, D. J. (1985). Exchangeability and related topics. In *École d'été de probabilités de saint-flour* (pp. 1–198). Springer.

Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, *15*(3), 147–149.

Ambartsumian, V. (1929). On the relationship between the solution and the resolvente of the integral equation of the radiative balance. *Zeitschrift für Physik*, *52*(3-4), 263–267.

Amsel, E., & Brock, S. (1996). The development of evidence evaluation skills. *Cognitive Development*, *11*(4), 523–550.

Anderson, J. (1990). *The adaptive character of thought.* Erlbaum.

Antoniak, C. E. (1974). Mixtures of dirichlet processes with applications to Bayesian nonparametric problems. *The Annals of Statistics*, 1152–1174.

Austerweil, J. L., & Griffiths, T. L. (2011). Seeking confirmation is rational for deterministic hypotheses. *Cognitive Science*, *35*(3), 499–526.

Baars, B. J. (1997). In the theatre of consciousness. global workspace theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies*, *4*(4), 292–309.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559.

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, *20*(3), 191–208.

Barber, D. (2012). *Bayesian reasoning and machine learning*. Cambridge University Press.

Baron, J. (2005). *Rationality and intelligence*. Cambridge University Press.

Baron, J., Beattie, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning: II. Congruence, information, and certainty. *Organizational Behavior and Human Decision Processes*, *42*(1), 88–110.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327–18332.

Battaglia, P. W., Pascanu, R., Lai, M., Rezende, D. J., et al. (2016). Interaction networks for learning about objects, relations and physics. In *Advances in neural information processing systems* (pp. 4502–4510).

Bayes, T., & Price. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions*, *53*, 370–418.

Bechlivanidis, C. (2015). The arrow of time through the causal lens: When causal beliefs determine temporal order. *Unpublished PhD thesis*.

Bechlivanidis, C., & Lagnado, D. A. (2013). Does the "why" tell us the "when"? *Psychological Science*, *24*(8), 1563–1572.

Bechlivanidis, C., & Lagnado, D. A. (2016). Time reordered: Causal perception guides the interpretation of temporal order. *Cognition*, *146*, 58–66.

Bellman, R. E. (2015). *Adaptive control processes: a guided tour*. Princeton University Press.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer, New York.

Blackmore, S. (2000). *The meme machine*. Oxford Paperbacks.

Bonawitz, E. B., Denison, S., Gopnik, A., & Griffiths, T. L. (2014). Win-stay, lose-sample: A simple sequential algorithm for approximating Bayesian inference. *Cognitive Psychology*, *74*, 35–65.

Bonawitz, E. B., van Schijndel, T. J., Friel, D., & Schulz, L. E. (2012). Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology*, *64*(4), 215–234.

Bramley, N. R., Dayan, P., Griffiths, T. L., & Lagnado, D. A. (2017). Formalizing Neurath's ship: Approximate algorithms for online causal learning. *Psychological Review*, to appear.

Bramley, N. R., Dayan, P., & Lagnado, D. A. (2015). Staying afloat on Neurath's boat: Heuristics for sequential causal learning. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 262–267). Austin, TX: Cognitive Science Society.

Bramley, N. R., Gerstenberg, T., & Lagnado, D. A. (2014). The order of things: Inferring causal structure from temporal patterns. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 236–242). Austin, TX: Cognitive Science Society.

Bramley, N. R., Gerstenberg, T., Mayrhofer, R., & Lagnado, D. A. (submitted). The role of time in causal learning.

Bramley, N. R., Gerstenberg, T., & Tenenbaum, J. B. (2016). Natural science: Active learning in dynamic physical microworlds. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 2567–2573). Austin, TX: Cognitive Science Society.

Bramley, N. R., Lagnado, D. A., & Speekenbrink, M. (2015). Conservative forgetful scholars: How people learn causal structure through interventions. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *41*(3), 708–731.

Bramley, N. R., Mayrhofer, R., Gerstenberg, T., & Lagnado, D. A. (2017). Causal learning from interventions and dynamics in continuous time. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Bramley, N. R., Nelson, J. D., Speekenbrink, M., Crupi, V., & Lagnado, D. A. (2014). *What should causal learners value?* Poster presented at the Annual Meeting of the Psychonomic Society.

Brass, M., & Haggard, P. (2007). To do or not to do: the neural signature of self-control. *The Journal of Neuroscience*, *27*(34), 9141–9145.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, *81*(3), 211–241.

Brent, R. P. (2013). *Algorithms for minimization without derivatives.* Courier Corporation.

Brighton, H., & Gigerenzer, G. (2012). Are rational actor models "rational" outside small worlds. *Evolution and Rationality: Decisions, Co-operation, and Strategic Behavior*, 84–109.

Broadbent, D. (1970). Stimulus set and response set: Two kinds of selective attention. In D. Mostofsky (Ed.), *Attention: Contemporary theory and analysis.* Appleton Century Crofts.

Brown, S. D., & Steyvers, M. (2009). Detecting and predicting changes. *Cognitive Psychology*, *58*(1), 49–67.

Bruner, J. S., Jolly, A., & Sylva, K. (1976). *Play: Its role in development and evolution.* Penguin.

Buchanan, D. W., Tenenbaum, J. B., & Sobel, D. M. (2010). Edge replacement and nonindependence in causation. In *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (pp. 919–924). Austin, TX: Cognitive Science Society.

Buehner, M. J., Cheng, P. W., & Clifford, D. (2003). From covariation to causation: a test of the assumption of causal power. *Journal of Experimental Psychology: Learning Memory & Cognition*, *29*, 1119–40.

Buehner, M. J., & Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychological Science*, *20*(10), 1221–1228.

Buehner, M. J., & May, J. (2002). Knowledge mediates the timeframe of covariation assessment in human causal induction. *Thinking & Reasoning*, *8*(4), 269–295.

Buehner, M. J., & May, J. (2003). Rethinking temporal contiguity and the judgement of causality: Effects of prior knowledge, experience, and reinforcement procedure. *The Quarterly Journal of Experimental Psychology Section A*, *56*(5), 865–890.

Buehner, M. J., & May, J. (2004). Abolishing the effect of reinforcement delay on human causal learning. *Quarterly Journal of Experimental Psychology Section B*, *57*(2), 179–191.

Buehner, M. J., & McGregor, S. (2006). Temporal delays can facilitate causal attribution: Towards a general timeframe bias in causal induction. *Thinking & Reasoning*, *12*(4), 353–378.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*(1), 3–5.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209–254).

Burns, P., & McCormack, T. (2009). Temporal information and children's and adults' causal inferences. *Thinking & Reasoning*, *15*(2), 167–196.

Butko, N. J., & Movellan, J. R. (2008). I-POMDP: An infomax model of eye movement. *The 7th IEEE International Conference on Development and Learning*(1), 139–144.

Carnap, R. (1962). *Logical foundations of probability.* University of Chicago press Chicago.

Cartwright, N. (1989). *Nature's capacities and their measurement.* Oxford: Clarendon.

Cartwright, N. (2001). What is wrong with Bayes nets? *The Monist*, *84*(2), 242–264.

Cartwright, N. (2007). *Hunting causes and using them*. Cambridge University Press Cambridge.

Casini, L., Illari, P. M., Russo, F., & Williamson, J. (2011). Models for prediction, explanation and control: recursive Bayesian networks. *Theoria*, *26*(1), 5–33.

Chang, M. B., Ullman, T., Torralba, A., & Tenenbaum, J. B. (2016). A compositional object-based approach to learning physical dynamics. *arXiv preprint arXiv:1612.00341*.

Chater, N., & Oaksford, M. (1993). Logicism, mental models and everyday reasoning. *Mind & Language*, *8*, 72–89.

Chater, N., & Oaksford, M. (1999). Ten years of the rational analysis of cognition. *Trends in Cognitive Sciences*, *3*(2), 57–65.

Chater, N., Tenenbaum, J. B., & Yuille, A. (2006). Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Sciences*, *10*(7), 288–291.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, *70*(5), 1098–1120.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*(2), 367–405.

Chib, S., & Jeliazkov, I. (2001). Marginal likelihood from the metropolis–hastings output. *Journal of the American Statistical Association*, *96*(453), 270–281.

Chopin, N. (2002). A sequential particle filter method for static models. *Biometrika*, *89*(3), 539–552.

Chow, Y. S., Robbins, H., & Siegmund, D. (1971). *Great expectations: The theory of optimal stopping*. Houghton Mifflin.

Christian, B., & Griffiths, T. (2016). *Algorithms to live by: The computer science of human decisions*. Macmillan.

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, 62.

Clarke, B., Leuridan, B., & Williamson, J. (2013). Modelling mechanisms with causal cycles. *Synthese*, 1–31.

Coenen, A., Rehder, R., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology*, *79*, 102–133.

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition*, *120*(3), 341–349.

Cooper, G. F. (1990). The computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence*, *42*(2), 393–405.

Cooper, G. F., & Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, *9*(4), 309–347.

Cooper, R. (2000). Simple heuristics could make us smart; but which heuristics do we apply when? *Behavioral and Brain Sciences*, *23*(746), 746.

Corner, A., Harris, A. J. L., & Hahn, U. (2010). Conservatism in belief revision and participant skepticism. In *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (Vol. 1625, p. 1630). Austin, TX: Cognitive Science Society.

Courville, A. C., & Daw, N. D. (2007). The rat as particle filter. In *Advances in Neural Information Processing Systems* (pp. 369–376).

Cover, T. M., & Thomas, J. A. (1991). Entropy, relative entropy and mutual information. *Elements of Information Theory*, *2*, 1–55.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.

Craver, C. F. (2006). When mechanistic models explain. *Synthese*, *153*(3), 355–376.

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating amazon's mechanical turk as a tool for experimental behavioral research. *PloS One*, *8*(3), e57410.

Danks, D. (2014). *Unifying the mind: Cognitive representations as graphical models.* MIT Press.

Dasgupta, I., Schulz, E., & Gershman, S. J. (2016). Where do hypotheses come from? *Center for Brains, Minds and Machines (preprint)*.

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*(2), 199–204.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879.

Dean, T., & Kanazawa, K. (1989). A model for reasoning about persistence and causation. *Computational Intelligence*, *5*(2), 142–150.

Dean Jr, D., & Kuhn, D. (2007). Direct instruction vs. discovery: The long view. *Science Education*, *91*(3), 384–397.

DeCarlo, L. T. (1992). Intertrial interval and sequential effects in magnitude scaling. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1080.

Denil, M., Agrawal, P., Kulkarni, T. D., Erez, T., Battaglia, P., & de Freitas, N. (2017).

Learning to perform physics experiments via deep reinforcement learning. *International Conference on Learning Representations*.

Dennett, D. C. (1988). The intentional stance in theory and practice. In R. Byrne & A. Whiten (Eds.), *Machiavellian intelligence* (pp. 180–202). Oxford, UK: Oxford University Press.

Dennett, D. C. (1991). *Consciousness explained.* London, UK: Penguin.

Deverett, B., & Kemp, C. (2012). Learning deterministic causal networks from observational data. In *Proceedings of the 34th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Dobson, A. J. (2010). *An introduction to generalized linear models.* CRC press.

Duhem, P. M. M. (1991). *The aim and structure of physical theory.* Princeton University Press.

Eberhardt, F., Glymour, C., & Scheines, R. (2012). On the number of experiments sufficient and in the worst case necessary to identify all causal relations among n variables. *arXiv preprint arXiv:1207.1389*.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 17–52).

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, *99*(1), 3.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis.* Cambridge, MA: MIT Press.

Estes, W. K. (1948). Discriminative conditioning: Effects of a pavlovian conditioned stimulus upon a subsequently established operant response. *Journal of Experimental Psychology*, *38*(2), 173.

Evans, J. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459.

Feldbaum, A. (1960). Dual control theory i–iv. *Avtomatika i Telemekhanika*.

Fenton, N., Neil, M., Lagnado, D., Marsh, W., Yet, B., & Constantinou, A. (2016). How to model mutually exclusive events based on independent causal pathways in bayesian network models. *Knowledge-based Systems*, *113*, 39–50.

Fernbach, P. M., Darlow, A., & Sloman, S. A. (2010). Neglect of alternative causes in predictive but not diagnostic reasoning. *Psychological Science*, *21*(3), 329–336.

Fernbach, P. M., Darlow, A., & Sloman, S. A. (2011). Asymmetries in predictive and diagnostic reasoning. *Journal of Experimental Psychology: General*, *140*(2), 168.

Fernbach, P. M., & Sloman, S. A. (2009). Causal learning with local computations.

Journal of Experimental Psychology: Learning, Memory & Cognition, *35*(3), 678.

Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review*, *90*(3), 239.

Frosch, C. A., McCormack, T., Lagnado, D. A., & Burns, P. (2012). Are Causal Structure and Intervention Judgments Inextricably Linked? A Developmental Study. *Cognitive Science*, *36*, 261–285.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis. texts in statistical science series.*

Geman, S., & Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the Bayesian restoration of images. *Pattern Analysis and Machine Intelligence*(6), 721–741.

Georges, A., Kotliar, G., Krauth, W., & Rozenberg, M. J. (1996). Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Reviews of Modern Physics*, *68*(1), 13.

Gershman, S. J. (2016). On the blessing of abstraction. *The Quarterly Journal of Experimental Psychology*, 1–5.

Gershman, S. J., Jones, C. E., Norman, K. A., Monfils, M.-H., & Niv, Y. (2013). Gradual extinction prevents the return of fear: implications for the discovery of state. *Frontiers in Behavioral Neuroscience*, *7*, 164.

Gershman, S. J., & Niv, Y. (2012). Exploring a latent cause theory of classical conditioning. *Learning & Behavior*, *40*(3), 255–268.

Gershman, S. J., Vul, E., & Tenenbaum, J. B. (2012). Multistability and perceptual inference. *Neural Computation*, *24*(1), 1–24.

Gerstenberg, T., Bechlivanidis, C., & Lagnado, D. A. (2013). Back on track: Backtracking in counterfactual reasoning. In *Proceedings of the 35th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2015). How, whether, why: Causal judgments as counterfactual contrasts. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 782–787). Austin, TX: Cognitive Science Society.

Gerstenberg, T., & Tenenbaum, J. B. (to appear). Intuitive theories. In M. Waldman (Ed.), *Oxford handbook of causal reasoning.* Oxford University Press.

Ghahramani, Z., & Griffiths, T. L. (2005). Infinite latent feature models and the indian buffet process. In *Advances in Neural Information Processing Systems* (pp. 475–482).

Gigerenzer, G. (2001). The adaptive toolbox. In G. Gigerenzer & R. Selten (Eds.),

Bounded rationality (pp. 37–50). MIT Press.

Gigerenzer, G., Todd, P. M., & ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Gilden, D. L. (2001). Cognitive emissions of 1/f noise. *Psychological Review*, *108*(1), 33.

Gleick, J. (1997). *Chaos: Making a new science*. Open Road Media.

Glymour, C. (1998). Learning causes: Psychological explanations of causal explanation. *Minds and Machines*, *8*(1), 39–60.

Glymour, C. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology*. Cambridge, MA: MIT Press.

Good, I. J. (1950). *Probability and the weighing of evidence*. Griffin, London.

Goodman, N. (1955). *Fact, fiction and forecast*. Harvard University Press.

Goodman, N. D. (2013). The principles and practice of probabilistic programming. *ACM SIGPLAN Notices*, *48*(1), 399–402.

Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A rational analysis of rule-based concept learning. *Cognitive Science*, *32*(1), 108–154.

Goodman, N. D., Tenenbaum, J. B., & Gerstenberg, T. (2015). Concepts in a probabilistic language of thought. In E. Margolis & S. Lawrence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 623–653). MIT Press.

Goodman, N. D., Ullman, T. D., & Tenenbaum, J. B. (2011). Learning a theory of causality. *Psychological Review*, *118*(1), 110–9.

Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. *Science*, *337*(6102), 1623–1627.

Gopnik, A., Glymour, C., Sobel, D., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*, 1–31.

Gopnik, A., & Sobel, D. M. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child Development*, *71*(5), 1205–1222.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, *37*(5), 620.

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, *138*(6), 1085.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? a comparative analysis of six preconceptions about internet questionnaires. *American Psychologist*, *59*(2), 93.

Goudie, R. J., & Mukherjee, S. (2011). *An efficient Gibbs sampler for structural inference in Bayesian networks* (Tech. Rep.). Centre for Research in Statistical Methodology, Warwick.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 424–438.

Greville, W. J., & Buehner, M. J. (2007). The influence of temporal distributions on causal induction from tabular data. *Memory & Cognition*, *35*(3), 444–453.

Greville, W. J., & Buehner, M. J. (2010). Temporal predictability facilitates causal learning. *Journal of Experimental Psychology: General*, *139*(4), 756–771.

Greville, W. J., Cassar, A. A., Johansen, M. K., & Buehner, M. J. (2013). Structural awareness mitigates the effect of delay in human causal learning. *Memory & Cognition*, 1–13.

Grice, G. R. (1948). The relation of secondary reinforcement to delayed reward in visual discrimination learning. *Journal of Experimental Psychology*, *38*(1), 1.

Griffiths, T. L. (2005). Causes, coincidences, and theories. *Unpublished doctoral dissertation*.

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, *7*, 217–229.

Griffiths, T. L., Lucas, C., Williams, J., & Kalish, M. L. (2009). Modeling human function learning with Gaussian processes. In *Advances in Neural Information Processing Systems* (pp. 553–560).

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*(4), 334–384.

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review*, *116*, 661–716.

Guez, A. (2015). Sample-based Search Methods for Bayes-Adaptive Planning. *Unpublished PhD thesis*.

Guez, A., Silver, D., & Dayan, P. (2012). Efficient bayes-adaptive reinforcement learning using sample-based search. In *Advances in Neural Information Processing Systems* (pp. 1025–1033).

Guez, A., Silver, D., & Dayan, P. (2013). Scalable and efficient bayes-adaptive reinforcement learning based on monte-carlo tree search. *Journal of Artificial Intelligence*

Research, *48*, 841–883.

Guo, X., Singh, S., Lee, H., Lewis, R. L., & Wang, X. (2014). Deep learning for real-time atari game play using offline monte-carlo tree search planning. In *Advances in Neural Information Processing Systems* (pp. 3338–3346).

Gureckis, T. M., & Markant, D. (2009). Active learning strategies in a spatial concept learning game. In *Proceedings of the 31st Annual Meeting of the Cognitive Science Society* (pp. 3145–3150). Austin, TX: Cognitive Science Society.

Gureckis, T. M., & Markant, D. B. (2012, September). Self-Directed Learning: A Cognitive and Computational Perspective. *Perspectives on Psychological Science*, *7*(5), 464–481.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382–385.

Hagmayer, Y., & Waldmann, M. R. (2002). How temporal assumptions influence causal judgments. *Memory & Cognition*, *30*(7), 1128–1137.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, *21*(06), 803–831.

Halpern, J. Y. (2016). *Actual causality*. MIT Press.

Halpern, J. Y., & Hitchcock, C. (2011). Actual causation and the art of modeling. *arXiv preprint arXiv:1106.2652*.

Halpern, J. Y., & Pearl, J. (2005). Causes and explanations: A structural-model approach. part i: Causes. *The British Journal for the Philosophy of Science*, *56*(4), 843–887.

Hamrick, J. B., Battaglia, P. W., Griffiths, T. L., & Tenenbaum, J. B. (2016). Inferring mass in complex scenes by mental simulation. *Cognition*, *157*, 61–76.

Hamrick, J. B., Smith, K. A., Griffiths, T. L., & Vul, E. (2015). Think again? the amount of mental simulation tracks uncertainty in the outcome. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Harman, G. (1986). *Change in view: Principles of reasoning.* Cambridge, MA: MIT Press.

Hartigan, J. A. (2012). *Bayes theory*. Springer Science & Business Media.

Hartigan, J. A., & Hartigan, P. (1985). The dip test of unimodality. *The Annals of Statistics*, 70–84.

Haugeland, J. (1989). *Artificial intelligence: The very idea*. MIT press.

Hauser, D. J., & Schwarz, N. (2015). Attentive turkers: Mturk participants perform

better on online attention checks than do subject pool participants. *Behavior research methods*, 1–8.

Heckerman, D., Geiger, D., & Chickering, D. M. (1995). Learning bayesian networks: The combination of knowledge and statistical data. *Machine learning*, *20*(3), 197–243.

Hofstadter, D. R. (1980). *Godel, Escher, Bach: An eternal golden braid*. New Society.

Hofstadter, D. R. (2007). *I am a strange loop*. Basic books.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, *24*(1), 1–55.

Hohwy, J. (2013). *The predictive mind*. Oxford University Press.

Holyoak, K. J., & Cheng, P. W. (2011, January). Causal learning and inference as a rational process: the new synthesis. *Annual Review of Psychology*, *62*, 135–63.

Hume, D. (1740). *A treatise of human nature*. Oxford Philosophical Texts (2000 reprint).

Hume, D. (1748). *An enquiry concerning human understanding*. Oxford University Press.

Hwang, C.-R. (1988). Simulated annealing: theory and applications. *Acta Applicandae Mathematicae*, *12*(1), 108–111.

Hyafil, L., & Rivest, R. L. (1976). Constructing optimal binary decision trees is np-complete. *Information Processing Letters*, *5*(1), 15–17.

Illari, P. M., & Williamson, J. (2012). What is a mechanism? thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, *2*(1), 119–135.

Jaakkola, T. S. (2001). A tutorial on variational approximation methods. In M. Opper (Ed.), *Advanced mean field methods: theory and practice* (p. 129). MIT Press.

James, W. (1890). *The principles of psychology*. Dover (2000 reprint).

Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs: General and Applied*, *79*(1), 1.

Jocham, G., Brodersen, K. H., Constantinescu, A. O., Kahn, M. C., Ianni, A. M., Walton, M. E., ... Behrens, T. E. (2016). Reward-guided learning with and without causal attribution. *Neuron*.

Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science*, *310*(5745), 116–119.

Jones, M., & Love, B. C. (2011). Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *The Behavioral and Brain Sciences*, *34*(4), 169–88.

Julier, S. J., & Uhlmann, J. K. (1997). New extension of the Kalman filter to nonlinear

systems. In *Proceedings of the 1997 AeroSense Conference on Photonic Quantum Computing* (pp. 182–193).

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: encoding time in neural network states. *Neuron*, *53*(3), 427–438.

Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology*, *57*, 227.

Kemp, C., Goodman, N. D., & Tenenbaum, J. B. (2010). Learning to learn causal models. *Cognitive Science*, *34*(7), 1185–243.

Kemp, C., & Tenenbaum, J. B. (2009). Structured statistical models of inductive reasoning. *Psychological Review*, *116*(1), 20.

Kemp, C., Tenenbaum, J. B., Niyogi, S., & Griffiths, T. L. (2010). A probabilistic model of theory formation. *Cognition*, *114*(2), 165–196.

Kim, N. S., & Ahn, W.-k. (2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology: General*, *131*(4), 451.

Kingman, J. F. C. (1993). *Poisson processes*. Wiley.

Kipling, R. (1902). *Just so stories*. Random House.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, *12*(1), 1–48.

Klahr, D., Fay, A. L., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology*, *25*(1), 111–146.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction effects of direct instruction and discovery learning. *Psychological science*, *15*(10), 661–667.

Klayman, J. (1995). Varieties of confirmation bias. *Psychology of Learning and Motivation*, *32*, 385–418.

Klayman, J., & Ha, Y.-w. (1989). Hypothesis testing in rule discovery: Strategy, structure, and content. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *15*(4), 596.

Klenske, E. D., & Hennig, P. (2016). Dual control for approximate bayesian reinforcement learning. *Journal of Machine Learning Research*, *17*(127), 1–30.

Koerber, S., Sodian, B., Thoermer, C., & Nett, U. (2005). Scientific reasoning in young children: Preschoolers' ability to evaluate covariation evidence. *Swiss Journal of Psychology*, *64*(3), 141–152.

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*(6971), 244–247.

Krippendorff, K. (2012). *Content analysis: An introduction to its methodology.* Sage.

Krynski, T. R., Tenenbaum, J. B., et al. (2007). The role of causality in judgment under uncertainty. *Journal of Experimental Psychology General*, *136*(3), 430.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, *96*(4), 674.

Kuhn, D., Amsel, E., O'Loughlin, M., Schauble, L., Leadbeater, B., & Yotive, W. (1988). *The development of scientific thinking skills.* Academic Press.

Kuhn, D., & Dean, D. (2005). Is developing scientific thinking all about learning to control variables? *Psychological Science*, *16*(11), 866–870.

Kuhn, G., Caffaratti, H. A., Teszka, R., & Rensink, R. A. (2014). A psychologically-based taxonomy of misdirection. *Frontiers in Psychology*, *5*, 1392.

Kuhn, T. S. (1962). *The structure of scientific revolutions.* University of Chicago Press.

Kushner, H. J. (1967). Dynamical equations for optimal nonlinear filtering. *Journal of Differential Equations*, *3*(2), 179–190.

Kushnir, T., Gopnik, A., Lucas, C., & Schulz, L. E. (2010). Inferring hidden causal structure. *Cognitive Science*, *34*(1), 148–60.

Lagnado, D. A. (2011). Thinking about evidence. In *Proceedings of the British Academy* (Vol. 171, pp. 183–223).

Lagnado, D. A., & Gerstenberg, T. (in press). Causation in legal and moral reasoning. In M. R. Waldmann (Ed.), *Oxford handbook of causal reasoning.* Oxford: Oxford University Press.

Lagnado, D. A., Gerstenberg, T., & Zultan, R. (2013). Causal responsibility and counterfactuals. *Cognitive Science*, *37*(6), 1036–1073.

Lagnado, D. A., & Sloman, S. A. (2002). Learning causal structure. In *Proceedings of the 24th Annual Meeting of the Cognitive Science Society.* Erlbaum.

Lagnado, D. A., & Sloman, S. A. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *30*, 856–876.

Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *32*(3), 451–60.

Lagnado, D. A., & Speekenbrink, M. (2010). The influence of delays in real-time causal learning. *The Open Psychology Journal*, *3*(2), 184–195.

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation: cues to causal structure. In A. Gopnik & L. E. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 154–72). London: Oxford

University Press.

Lakatos, I. (1976). Falsification and the methodology of scientific research programmes. In *Can theories be refuted?* (pp. 205–259). Springer.

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, *350*(6266), 1332–1338.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (to appear). Building machines that learn and think like people. *Behavioral and Brain Sciences*.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159–174.

Lauritzen, S. L., Dawid, A. P., Larsen, B. N., & Leimer, H.-G. (1990). Independence properties of directed markov fields. *Networks*, *20*(5), 491–505.

Lauritzen, S. L., & Richardson, T. S. (2002). Chain graph models and their causal interpretations. *Journal of the Royal Statistical Society*, *64*(3), 321–348.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Lee, H. S., & Holyoak, K. J. (2008). The role of causal models in analogical inference. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *34*(5), 1111.

Lerer, A., Gross, S., & Fergus, R. (2016). Learning physical intuition of block towers by example. *arXiv preprint arXiv:1603.01312*.

Levy, R. P., Reali, F., & Griffiths, T. L. (2009). Modeling the effects of memory on human online sentence processing with particle filters. In *Advances in Neural Information Processing Systems* (pp. 937–944).

Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice*. Sage.

Lieder, F., Griffiths, T., & Goodman, N. (2012). Burn-in, bias, and the rationality of anchoring. In *Advances in Neural Information Processing Systems* (pp. 2690–2798).

Lieder, F., & Griffiths, T. L. (2015). When to use which heuristic: A rational solution to the strategy selection problem. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1362–1367). Austin, TX: Cognitive Science Society.

Lieder, F., Griffiths, T. L., Huys, Q. J. M., & Goodman, N. D. (under review). Reinterpreting anchoring-and-adjustment as rational use of cognitive resource.

Lindley, D. V. (1956). On a measure of the information provided by an experiment.

The Annals of Mathematical Statistics, 986–1005.

Liu, J. S., & Chen, R. (1998). Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association*, *93*(443), 1032–1044.

London, J. (2012). *Hearing in time: Psychological aspects of musical meter.* Oxford: Oxford University Press.

Lopez, F. J., Cobos, P. L., Caño, A., & Shanks, D. R. (1998). The rational analysis of human causal and probability judgment. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (p. 314-352). Oxford: Oxford University Press.

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review*, *115*(4), 955.

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, *131*(2), 284–299.

Lucas, C. G., & Griffiths, T. L. (2010). Learning the form of causal relationships using hierarchical bayesian models. *Cognitive Science*, *34*(1), 113–147.

Lucas, C. G., Holstein, K., & Kemp, C. (2014). Discovering hidden causes using statistical evidence. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Luce, D. R. (1959). *Individual choice behavior.* New York: Wiley.

MacKay, D. J. C. (2003). *Information theory, inference and learning algorithms.* Cambridge University Press.

Mackintosh, N. J. (1983). *Conditioning and associative learning.* Clarendon Press Oxford.

Macready, W. G., & Wolpert, D. H. (1998). Bandit problems and the exploration/exploitation tradeoff. *IEEE Transactions on Evolutionary Computation*, *2*(1), 2–22.

Madigan, D., & Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using occam's window. *Journal of the American Statistical Association*, *89*(428), 1535–1546.

Madigan, D., York, J., & Allard, D. (1995). Bayesian graphical models for discrete data. *International Statistical Review*, 215–232.

Malthus, T. R. (1888). *An essay on the principle of population: or, a view of its past and present effects on human happiness.* Reeves & Turner.

Markant, D. B., & Gureckis, T. M. (2010). Category learning through active sampling. *Proceedings of the of the 32nd Annual Meeting of the Cognitive Science Society*, 248–253.

Markant, D. B., & Gureckis, T. M. (2012). Does the utility of information influence sampling behavior? In *Proceedings of the 34th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Markant, D. B., & Gureckis, T. M. (2014). Is it better to select or to receive? learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, *143*(1), 94.

Markant, D. B., Settles, B., & Gureckis, T. M. (2015). Self-directed learning favors local, rather than global, uncertainty. *Cognitive Science*.

Marr, D. (1982). *Vision*. New York: Freeman & Co.

Marshall, B. J., Armstrong, J. A., McGechie, D. B., & Glancy, R. J. (1985). Attempt to fulfil koch's postulates for pyloric campylobacter. *The Medical Journal of Australia*, *142*(8), 436–439.

Mason, W., & Suri, S. (2012). Conducting behavioral research on amazon's mechanical turk. *Behavior Research Methods*, *44*(1), 1–23.

Mayrhofer, R., Goodman, N. D., Waldmann, M. R., & Tenenbaum, J. B. (2008). Structured correlation from the causal background. In *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (pp. 303–308). Austin, TX: Cognitive Science Society.

Mayrhofer, R., & Waldmann, M. R. (2011). Heuristics in covariation-based induction of causal models: Sufficiency and necessity priors. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 3110–3115). Austin, TX: Cognitive Science Society.

Mayrhofer, R., & Waldmann, M. R. (2015). Agents and causes: Dispositional intuitions as a guide to causal structure. *Cognitive Science*, *39*(1), 65–95.

Mayrhofer, R., & Waldmann, M. R. (2016). Sufficiency and necessity assumptions in causal structure induction. *Cognitive Science*, *40*(8).

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naïve beliefs about the motion of objects. *Science*, *210*(4474), 1138–1141.

McCormack, T., Bramley, N. R., Frosch, C., Patrick, F., & Lagnado, D. A. (2016). Children's use of interventions to learn causal structure. *Journal of Experimental Child Psychology*, *141*, 1–22.

McCormack, T., Frosch, C., Patrick, F., & Lagnado, D. (2015). Temporal and statistical information in causal structure learning. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *41*(2), 395.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in

nervous activity. *The Bulletin of Mathematical Biophysics*, *5*(4), 115–133.

Meder, B., Gerstenberg, T., Hagmayer, Y., & Waldmann, M. R. (2010). Observing and intervening: Rational and heuristic models of causal decision making. *The Open Psychology Journal*, *3*, 119–135.

Meder, B., Mayrhofer, R., & Waldmann, M. R. (2014). Structure induction in diagnostic causal reasoning. *Psychological Review*, *121*(3), 277.

Meder, B., & Nelson, J. (2012). Information search with situation-specific reward functions. *Judgment and Decision Making*, *7*(2), 119–148.

Meier, K. M., & Blair, M. R. (2012). Waiting and weighting: Information sampling is a balance between efficiency and error-reduction. *Cognition*, *126*(2), 319–325.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, *21*(6), 1087–1092.

Michotte, A. (1946/1963). *The perception of causality.* Basic Books.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81.

Miller, G. A. (1984). Informavores. *The study of information: Interdisciplinary messages*.

Miyazaki, M., Nozaki, D., & Nakajima, Y. (2005). Testing bayesian models of human coincidence timing. *Journal of Neurophysiology*, *94*(1), 395–399.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533.

Murphy, K. P. (2001). *Active learning of causal Bayes net structure* (Tech. Rep.). UC Berkeley.

Myers, J. L. (2014). Probability learning and sequence learning. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (pp. 171–205).

Navarro, D. J., & Perfors, A. F. (2011). Hypothesis generation, sparse categories, and the positive test strategy. *Psychological Review*, *118*(1), 120.

Nelson, J. D. (2005). Finding useful questions: on Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, *112*(4), 979–99.

Nelson, J. D., Divjak, B., Gudmundsdottir, G., Martignon, L. F., & Meder, B. (2014). Children's sequential information search is sensitive to environmental probabilities. *Cognition*, *130*(1), 74–80.

Nelson, J. D., McKenzie, C. R. M., Cottrell, G. W., & Sejnowski, T. J. (2010). Experience Matters: Information Acquisition Optimizes Probability Gain. *Psychological*

Science, *21*(7), 960–969.

Neurath, O. (1932). Protocol statements. In *Philosophical papers 1913–1946* (pp. 91–99). Springer.

Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 104) (No. 9). Prentice-Hall Englewood Cliffs, NJ.

Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing "one-reason" decision making. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *29*(1), 53–65.

Newell, B. R., & Shanks, D. R. (2004). On the role of recognition in decision making. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *30*(4), 923–35.

Newell, B. R., & Shanks, D. R. (2007). Perspectives on the "tools" of decision-making. In M. J. Roberts (Ed.), *Integrating the mind: Domain general versus domain specific processes in higher cognition* (pp. 31–151). Hove, UK: Psychology Press.

Newell, B. R., & Shanks, D. R. (2014). Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences*, *37*(01), 1–19.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175.

Nielsen, F., & Nock, R. (2011). A closed-form expression for the sharma–mittal entropy of exponential families. *Journal of Physics A: Mathematical and Theoretical*, *45*(3), 032003.

Nikolic, M., & Lagnado, D. A. (2015). There aren't plenty more fish in the sea: A causal network approach. *British Journal of Psychology*, *106*(4), 564–582.

Nodelman, U., Shelton, C. R., & Koller, D. (2002). Continuous time bayesian networks. In *Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence* (pp. 378–387).

Norberg, T. (1986). Random capacities and their distributions. *Probability theory and related fields*, *73*(2), 281–297.

Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-regulation* (pp. 1–18). Springer.

Nyberg, E., & Korb, K. (2006). Informative interventions. In F. Russo & J. Williamson (Eds.), *Causality and probability in the sciences.* College: London.

Oaksford, M., & Chater, N. (1994). A rational analysis ofthe selection task as optimal data selection. *Psychological Review*, *101*(4), 608–631.

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

Oppenheimer, D. M. (2004). Spontaneous discounting of availability in frequency judgment tasks. *Psychological Science*, *15*(2), 100–105.

Oppenheimer, D. M., Tenenbaum, J. B., & Krynski, T. R. (2013). Categorization as causal explanation: Discounting and augmenting in a bayesian framework. *Psychology of Learning and Motivation*, *58*, 203–231.

Osman, M. (2011). *Controlling uncertainty: Decision making and learning in complex worlds*. Wiley.

Pacer, M. D., & Griffiths, L. (2012). Elements of a rational framework for continuous-time causal induction. In *Proceedings of the 34th Annual Meeting of the Cognitive Science Society* (Vol. 1, pp. 833–838). Austin, TX: Cognitive Science Society.

Pacer, M. D., & Griffiths, T. L. (2011). A rational model of causal induction with continuous causes. In *Advances in Neural Information Processing Systems* (pp. 2384–2392).

Pacer, M. D., & Griffiths, T. L. (2015). Upsetting the contingency table: Causal induction over sequences of point events. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Park, J., & Sloman, S. A. (2013). Mechanistic beliefs determine adherence to the markov property in causal reasoning. *Cognitive Psychology*, *67*(4), 186–216.

Parpart, P., Jones, M., & Love, B. (in revision). Heuristics as Bayesian inference.

Patil, K. R., Zhu, X., Kopeć, Ł., & Love, B. C. (2014). Optimal teaching for limited-capacity human learners. In *Advances in Neural Information Processing Systems* (pp. 2465–2473).

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Francisco, CA: Morgan Kaufmann.

Pearl, J. (2000). *Causality*. New York: Cambridge University Press (2nd edition).

Peirce, C. S. (1955). Abduction and induction. *Philosophical writings of Peirce*, *11*.

Perales, J. C., & Shanks, D. R. (2007). Models of covariation-based causal judgment: A review and synthesis. *Psychonomic Bulletin & Review*, *14*(4), 577–596.

Petrov, A. A., & Anderson, J. R. (2005). The dynamics of scaling: a memory-based anchor model of category rating and absolute identification. *Psychological Review*, *112*(2), 383.

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, *497*(7447), 74–79.

Piaget, J., & Valsiner, J. (1930). *The child's conception of physical causality*. Transaction Pub.

Piekny, J., & Maehler, C. (2013). Scientific reasoning in early and middle childhood:

The development of domain-general evidence evaluation, experimentation, and hypothesis generation skills. *British Journal of Developmental Psychology*, *31*(2), 153–179.

Pitman, J. (2002). *Combinatorial stochastic processes.* Lecture Notes for St. Flour Summer School.

Puterman, M. L. (2009). *Markov decision processes: discrete stochastic dynamic programming* (Vol. 414). Wiley.

Quine, W. v. O. (1969). *Word and object*. MIT press.

Raiffa, H. (1974). *Applied statistical decision theory.* Wiley.

Rehder, R. (2003). A causal-model theory of conceptual representation and categorization. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *29*(6), 1141.

Rehder, R. (2014). Independence and dependence in human causal reasoning. *Cognitive Psychology*, *72*, 54–107.

Rehder, R. (2016). Reasoning with causal cycles. *Cognitive Science*, *to appear*.

Rehder, R., & Hastie, R. (2001). Causal knowledge and categories: the effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General*, *130*(3), 323.

Reimers, S., & Stewart, N. (2015). Presentation and response timing accuracy in adobe flash and html5/javascript web experiments. *Behavior Research Methods*, *47*(2), 309–327.

Renyi, A. (1961). On measures of entropy and information. In *The fourth Berkeley symposium on mathematical statistics and probability* (pp. 547–561).

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, *2*, 64–99.

Rips, L. J. (2010). Two causal theories of counterfactual conditionals. *Cognitive Science*, *34*(2), 175–221.

Rips, L. J., & Edwards, B. J. (2013). Inference and explanation in counterfactual reasoning. *Cognitive Science*, *37*(6), 1107–1135.

Robinson, R. W. (1977). Counting unlabeled acyclic digraphs. In *Combinatorial mathematics v* (pp. 28–43). Springer.

Rothe, A., Deverett, B., Mayrhofer, R., & Kemp, C. (2016). *Successful structure learning from observational data.*

Rothe, A., Lake, B. M., & Gureckis, T. M. (2016). Asking and evaluating natural language questions. In *Proceedings of the 38th Annual Meeting of the Cognitive*

Science Society. Austin, TX: Cognitive Science Society.

Rottman, B. M., & Hastie, R. (2014). Reasoning about causal relationships: Inferences on causal networks. *Psychological bulletin*, *140*(1), 109.

Rottman, B. M., & Hastie, R. (2016, jun). Do people reason rationally about causally related events? Markov violations, weak inferences, and failures of explaining away. *Cognitive Psychology*, *87*, 88–134.

Rottman, B. M., & Keil, F. C. (2012). Causal structure learning over time: observations and interventions. *Cognitive psychology*, *64*(1), 93–125.

Ruggeri, A., & Lombrozo, T. (2014). Learning by asking: How children ask questions to achieve efficient search. In *36th annual meeting of the cognitive science society* (pp. 1335–1340). Austin, TX: Cognitive Science Society.

Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory & Cognition*, *17*(6), 759–769.

Safire, W. (1989). On Language: The Elision Fields. *New York Times*, 16–18.

Salmon, W. (1994). Causality without counterfactuals. *Philosophy of Science*, *61*(2), 297–312.

Sanborn, A. N. (2015). Types of approximation for probabilistic cognition: Sampling and variational. *Brain and cognition*, *to appear*.

Sanborn, A. N., & Chater, N. (2016). Bayesian brains without probabilities. *Trends in Cognitive Sciences*.

Sanborn, A. N., Griffiths, T. L., & Navarro, D. J. (2010). Rational approximations to rational models: alternative algorithms for category learning. *Psychological Review*, *117*(4), 1144.

Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and Newtonian mechanics for colliding objects. *Psychological Review*, *120*(2), 411.

Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology*, *49*(1), 31–57.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, *32*(1), 102.

Schlottmann, A. (1999). Seeing it happen and knowing how it works: How children understand the relation between perceptual causality and underlying mechanism. *Developmental Psychology*, *35*, 303–317.

Schulz, E., Konstantinidis, E., & Speekenbrink, M. (2016). Putting bandits into context: How function learning supports decision making. *bioRxiv*, 081091.

Schulz, L. E. (2001). Do-calculus?: Adults and preschoolers infer causal structure from patterns of outcomes following interventions. In *Second Biennial Meeting of the Cognitive Development Society.*

Schulz, L. E. (2012). The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in cognitive sciences*, *16*(7), 382–389.

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: preschoolers engage in more exploratory play when evidence is confounded. *Developmental psychology*, *43*(4), 1045.

Schulz, L. E., Goodman, N. D., Tenenbaum, J. B., & Jenkins, A. C. (2008). Going beyond the evidence: Abstract laws and preschoolers' responses to anomalous data. *Cognition*, *109*(2), 211–223.

Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental science*, *10*(3), 322–332.

Schulz, L. E., Kushnir, T., & Gopnik, A. (2007). Learning from doing: Intervention and causal inference. *Causal learning: Psychology, philosophy, and computation*, 67–85.

Schwartzman, H. (2012). *Transformations: The anthropology of children's play.* Springer Science & Business Media.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

Settles, B. (2012). Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *6*(1), 1–114.

Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. *The Psychology of Learning and Motivation*.

Shanks, D. R., & Lagnado, D. A. (2000). Sub-optimal reasons for rejecting optimality. *Behavioral and Brain Sciences*, *23*, 761–2.

Shanks, D. R., Tunney, R. J., & McCarthy, J. D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making*, *15*(3), 233–250.

Shannon, C. E. (1951). Prediction and entropy of printed english. *The Bell System Technical Journal*, *30*, 50–64.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: Ii. perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127.

Shultz, T. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development*, *47*(1), 1–51.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129–38.

Simon, H. A. (1982). *Models of bounded rationality: Empirically grounded economic reason*. MIT press.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*(7), 261–267.

Skinner, B. F. (1963). Operant behavior. *American Psychologist*, *18*(8), 503.

Skinner, B. F. (1990). *The behavior of organisms: An experimental analysis*. Skinner Foundation.

Sloman, S. A. (2005). *Causal models: How people think about the world and its alternatives*. Oxford University Press.

Sloman, S. A., & Lagnado, D. A. (2005, January). Do we "do"? *Cognitive Science*, *29*(1), 5–39.

Sloman, S. A., & Lagnado, D. A. (2015). Causality in thought. *Annual Review of Psychology*, *66*, 223–247.

Sloman, S. A., Love, B. C., & Ahn, W.-K. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, *22*(2), 189–228.

Smith, A. R. (1978). Color gamut transform pairs. *ACM Siggraph Computer Graphics*, *12*(3), 12–19.

Smith, K. A., & Vul, E. (2013). Sources of uncertainty in intuitive physics. *Topics in Cognitive Science*, *5*(1), 185–199.

Smith, K. A., & Vul, E. (2014). Looking forwards and backwards: Similarities and differences in prediction and retrodiction. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 1467–1472). Austin, TX: Cognitive Science Society.

Snow, J. (1855). *On the mode of communication of cholera*. John Churchill.

Sobel, D. M., & Kushnir, T. (2006). The importance of decision making in causal learning from interventions. *Memory & Cognition*, *34*(2), 411–419.

Sobel, D. M., & Sommerville, J. A. (2009). Rationales in children's causal learning from others' actions. *Cognitive Development*, *24*(1), 70–79.

Sobel, D. M., & Sommerville, J. A. (2010). The importance of discovery in children's causal learning from interventions. *Frontiers in Psychology*, *1*(76).

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and bayesian reasoning in preschoolers. *Cognitive Science*, *28*(3), 303–333.

Speekenbrink, M., & Konstantinidis, E. (2014). Uncertainty and exploration in a restless

bandit task. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Speekenbrink, M., & Shanks, D. R. (2010). Learning in a changing environment. *Journal of Experimental Psychology: General*, *139*(2), 266.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search.* Cambridge, MA: MIT Press (2nd edition, 2000).

Stewart, N. (2009). Decision by sampling: The role of the decision environment in risky choice. *The Quarterly Journal of Experimental Psychology*, *62*(6), 1041–1062.

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1–26.

Steyvers, M., Lee, M. D., & Wagenmakers, E.-J. (2009). A bayesian analysis of human decision-making on bandit problems. *Journal of Mathematical Psychology*, *53*(3), 168–179.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*, 453–489.

Strevens, M. (2008). *Depth: An account of scientific explanation.* Harvard University Press.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction.* Cambridge University Press.

Tenenbaum, J. B., Griffiths, T. L., & Niyogi, S. (2007). Intuitive theories as grammars for causal inference. *Causal learning: Psychology, philosophy, and computation*, 301–322.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*(6022), 1279–1285.

Thagard, P. (1992). *Conceptual revolutions.* Princeton University Press.

Treisman, M., & Williams, T. C. (1984). A theory of criterion setting with an application to sequential dependencies. *Psychological Review*, *91*(1), 68.

Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, *52*(1-2), 479–487.

Tsamardinos, I., Brown, L. E., & Aliferis, C. F. (2006). The max-min hill-climbing Bayesian network structure learning algorithm. *Machine Learning*, *65*(1), 31–78.

Tversky, A., & Edwards, W. (1966). Information versus reward in binary choices. *Journal of Experimental Psychology*, *71*(5), 680.

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141–162). Springer.

Ullman, T., Goodman, N., & Tenenbaum, J. B. (2012). Theory acquisition as stochastic

search in a language of thought. *Cognitive Development*, *27*, 455–480.

Ullman, T., Stuhlmüller, A., Goodman, N., & Tenenbaum, J. B. (2014). Learning physics from dynamical scenes. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

van Rooij, I., Wright, C. D., Kwisthout, J., & Wareham, T. (2014). Rational analysis, intractability, and the prospects of 'as if'-explanations. *Synthese*, 1–20.

Verschuur, G. L. (1996). *Hidden attraction: the history and mystery of magnetism.* Oxford University Press.

Vul, E., Goodman, N. D., Griffiths, T. L., & Tenenbaum, J. B. (2009). One and done? optimal decisions from very few samples. In *Proceedings of the 31st Annual Meeting of the Cognitive Science Society* (Vol. 1, pp. 66–72). Austin, TX: Cognitive Science Society.

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, *14*(1), 101–118.

Waldmann, M. R., Cheng, P., Hagmayer, Y., & Blaisdell, A. (2008). Causal learning in rats and humans: A minimal rational model. In M. Oaksford & N. Chater (Eds.), *The probabilistic mind* (pp. 453–484). Oxford: Oxford University Press.

Waldmann, M. R., & Hagmayer, Y. (2005). Seeing versus doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning Memory & Cognition*, *31*(2), 216–227.

Waldmann, M. R., Hagmayer, Y., & Blaisdell, A. P. (2006). Beyond the information given: Causal models of learning and reasoning. *Current Directions in Psychological Science*, *15*, 307–311.

Waldmann, M. R., & Mayrhofer, R. (2016). Hybrid causal representations. *Psychology of Learning & Motivation*, *65*, 85–127.

Walsh, C., & Sloman, S. A. (2008). Updating beliefs with causal models: Violations of screening off. In M. Gluck, J. R. Anderson, & S. M. Kosslyn (Eds.), *Memory and Mind: A festschrift for Gordon H. Bower* (pp. 345–58). Taylor & Francis.

Waskan, J. (2006). *Models and Cognition.* MIT Press.

Weber, E. H. (1834). *De pulsu, resorptione, auditu et tactu: Annotationes anatomicae et physiologicae.* Koehler.

Weierstrass, K. (1902). Mathematische werke iv. *Georg Olms*, *4*.

White, P. A. (1995). *The understanding of causation and the production of action: From infancy to adulthood.* Psychology Press.

White, P. A. (2006a). The causal asymmetry. *Psychological Review*, *113*(1), 132.

White, P. A. (2006b). How well is causal structure inferred from cooccurrence information? *European Journal of Cognitive Psychology*, *18*(3), 454–480.

White, P. A. (2009). Perception of forces exerted by objects in collision events. *Psychological Review*, *116*(3), 580.

Whitehead, A. N., & Russell, B. (1912). *Principia mathematica*. Rough Draft.

Williamson, J. (2004). Causality. In *Handbook of philosophical logic* (pp. 1–30).

Williamson, J., & Gabbay, D. (2005). Recursive causality in Bayesian networks and self-fibring networks. *Laws and Models in the Sciences*, 173–221.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, *55*, 235–269.

Wolfe, J. B. (1921). The effect of delayed reward upon learning in the white rat. *Journal of Comparative Psychology*, *17*(1), 1.

Wolff, P., & Barbey, A. K. (2015). Causal reasoning with forces. *Frontiers in Human Neuroscience*, *9*(1), 1–21.

Wolff, P., & Shepard, J. (2013). Causation, touch, and the perception of force. *Psychology of Learning and Motivation*, *58*, 167–202.

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE transactions on evolutionary computation*, *1*(1), 67–82.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press.

Woodward, J. (2011). Mechanisms revisited. *Synthese*, *183*(3), 409–427.

Wu, J., Yildirim, I., Lim, J. J., Freeman, B., & Tenenbaum, J. (2015). Galileo: Perceiving physical object properties by integrating a physics engine with deep learning. In *Advances in Neural Information Processing Systems* (pp. 127–135).

Yeung, S., & Griffiths, T. L. (2011). Estimating human priors on causal strength. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 1709–1714). Austin, TX: Cognitive Science Society.

Yeung, S., & Griffiths, T. L. (2015). Identifying expectations about the strength of causal relationships. *Cognitive Psychology*, *76*, 1–29.

Yu, A., & Dayan, P. (2003). Expected and unexpected uncertainty: ACh and NE in the neocortex. *Advances in Neural Information Processing Systems*, 173–180.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review*, *20*(1), 99–149.