Mechanisms of Active Causal Learning

Neil R. Bramley, David A. Lagnado and Maarten Speekenbrink.

University College London

Author Note

Neil R. Bramley, Department of Cognitive, Perceptual & Brain Sciences, University College London; David A. Lagnado, Department of Cognitive, Perceptual & Brain Sciences, University College London; Maarten Speekenbrink, Department of Cognitive, Perceptual & Brain Sciences, University College London.

Correspondence regarding this article should be sent to David Lagnado, Department of Cognitive, Perceptual & Brain Sciences, University College London, London, WC1E 6BT

E-mail: d.lagnado@ucl.ac.uk.

Tel: 44 (0) 20 7679 5389

Fax: 44 (0) 20 7436 4276

Word count: 11,732

Abstract

How well can people learn through active interventions on a causal system? Previous research has focused on simple systems and single interventions. This paper describes two experiments examining how people break down complex causal learning problems; how they learn without the support of contextual cues; and how contextual cues themselves are learned. Participants identified the causal structure underlying probabilistic patterns of data by the free selection of multiple interventions. Their interventions and model choices were measured against an optimal Bayesian model. Successful participants were systematic and efficient in their interventional learning, and several simple strategies were identified based on their fit to individual participants' actions and choices. We conclude that causal structure learning is likely to be based in simple action-selection and causal attribution mechanisms.

Keywords: Causal learning, Bayesian modelling; Heuristics; Hierarchical learning; Causal Bayesian networks; Multiple cue learning

Mechanisms of Active Causal Learning

Introduction

Causality in cognition

Causal beliefs are fundamental to how cognitive agents learn about and interact with an uncertain world (Pearl 2000; Sloman, 2005; Tenenbaum, Kemp, Griffiths & Goodman, 2011). Positing causal connections to explain patterns in observational data allows agents to form expectations about the consequences of their interactions with their environment, and to reason about the possible causes of new observations. This is because interpreting observations within a causal structure limits their implications to a manageable number of causally related items, facilitating and guiding inference. Moreover, the recent causal cognition literature supports the idea that our causal knowledge is represented in a structural and model-based way, rather than via simple associations (see for example Cheng, 1997; Lagnado et al., 2007; Shanks, 2007; Sloman, 2005; Waldmann, Hagmayer & Blaisdell, 2006; Woodward, 2003). Thus, an ability to identify and exploit stable causal structure in the environment is a central part of cognition.

Causal Bayesian networks

Causal Bayesian networks (CBNs, Pearl, 2000), have become a dominant method of formalising causal information. CBNs are based in directed acyclic graphs (see Figure 1) where *nodes* represent propositional variables of interest (such as whether or not a particular event occurs) and *directed edges* represent causal dependencies between those variables. These causal dependencies are quantified in the CBN by defining a conditional probability for each node, conditional on the states of its *parents* (any variables which have directed edges leading to that node). This is paired with the "causal Markov condition" (Pearl, 2000), which states that all variables are probabilistically independent of their non-descendants (other variables which are

not direct or indirect effects of that variable) given their parents. In this way a CBN can support inferences about particular propositional variables based on the states of only those variables that are *causally relevant* according to the graph. For example, in Figure 1 there is no direct effect of X_3 on X_5, but only an indirect effect mediated by X_4. Additionally, X_1, X_2 and X_3 are *screened off* from X_5 by X_4. This means that if you know the state of X_4, finding out about the states of any other nodes will tell you nothing more about the state of X_5. Rational inferences about any propositional variables are thus easily computed from the joint probability distribution implied by a CBN:

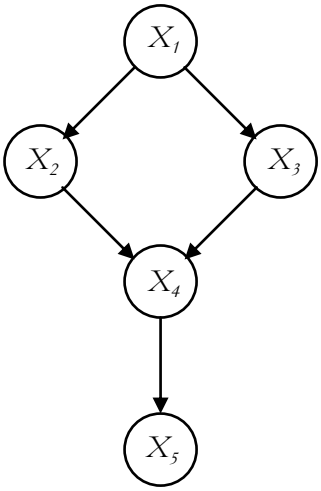

Figure 1. The directed acyclic graph for a simple 4 variable Bayes net. The arrows are the directed edges and circles are the nodes.

When an effect has multiple potential causes, as is the case with X_4, there are intuitively several ways in which they can combine. If the causes influence the effect independently, the probability of the effect is usually considered to be a *noisy-OR* function of the individual causal strengths the causes (Holyoak & Cheng, 2011). This is given by

$$P(E) = 1 - (1 - B)\prod_i (1 - p_i)^{Ci}$$

where the probability of the occurrence of an effect (E) given its *base rate* of occurrence (B) and *i* causes each with an independent probability of producing their effect (p_i), and a parameter C_i that takes a value of 1 if the cause is present and 0 if it is absent.

For example, suppose an abnormal x-ray result can be caused by cystic fibrosis or by tuberculosis. The probability of getting an abnormal x-ray result is thus given by:

$$P(xray) = 1 - (1 - P(xray|cystic\ fibrosis))^{C_1}(1 - P(xray|tuberculosis))^{C_2}$$

If we want to account for the fact that abnormal x-ray results are sometimes obtained for other reasons outside the scope of this model we can simply add a *base rate* of occurrence to this equation, giving:

$$P(xray) = 1 - (1 - B)(1 - P(xray|cystic\ fibrosis))^{C_1}(1 - P(xray|tuberculosis))^{C_2}$$

However, in many real world cases, it may not be the case that causes influence the effect independently. For example, two medications which individually raise one's probability of recovery might, in combination, lower the probability of recovery. However, it has been argued that without specific domain information people normally assume that multiple causes do not interact and are individually sufficient (Griffiths & Tenenbaum, 2009).

CBNs are powerful because they enable probabilistic inferences can take place when what you have to work with is a network of subjectively uncertain propositional beliefs about the world. Intuitively, observations of the states of particular propositional variables in the world lead us to adjust our subjective degrees of belief about the probability of various other causally relevant propositions, but not about beliefs which we think are causally unconnected or screened off by our causal model. Additionally, some noise is assumed by allowing that variables have base rates of occurrence due to exogenous causes (Cheng, 1997). These aspects make CBNs well suited to capturing approximate (partial but robust) structural interpretations of the kind of

complex, noisy data constitutive of experience (e.g. Glymour, 2001; Gopnik et al, 2004). In sum, the use of CBNs to model causal representation captures what is important about causal knowledge: its role in simplifying and facilitating reasoning from data (Sloman, 2005).

Intervention

The process of learning the causal structure of a system is commonly divided into two aspects. Passively *observing* the behaviour of the system allows learning about what patterns of (undirected) dependency it exhibits, while actively *intervening* to isolate and test subparts of it is necessary to identify its unique causal structure, at least without the help of additional cues (Lagnado & Sloman, 2002, 2004; Steyvers et al, 2003). The advantage of interventional learning is demonstrated by the fact that there are multiple ways of linking up and parameterising Bayes nets over sets of variables that are equally consistent with the same data. As a simple case, if two variables (A and B) covary this could either be due to A's causing B or vice-versa. This is the case regardless of how frequently either one activates with or without the other. The two possible structural interpretations just imply differing parameters of causal strength and base rate. Similarly, three variables (A, B and C) could exhibit the same patterns of dependence and independence if linked in a chain (A→B→C) or with B as a common cause (A←B→C). This is because, in both cases, each variable is probabilistically dependent on both of the other two, but A and C become independent of each other conditional on B. These observationally equivalent structures form "Markov equivalence classes" (Figure 2).

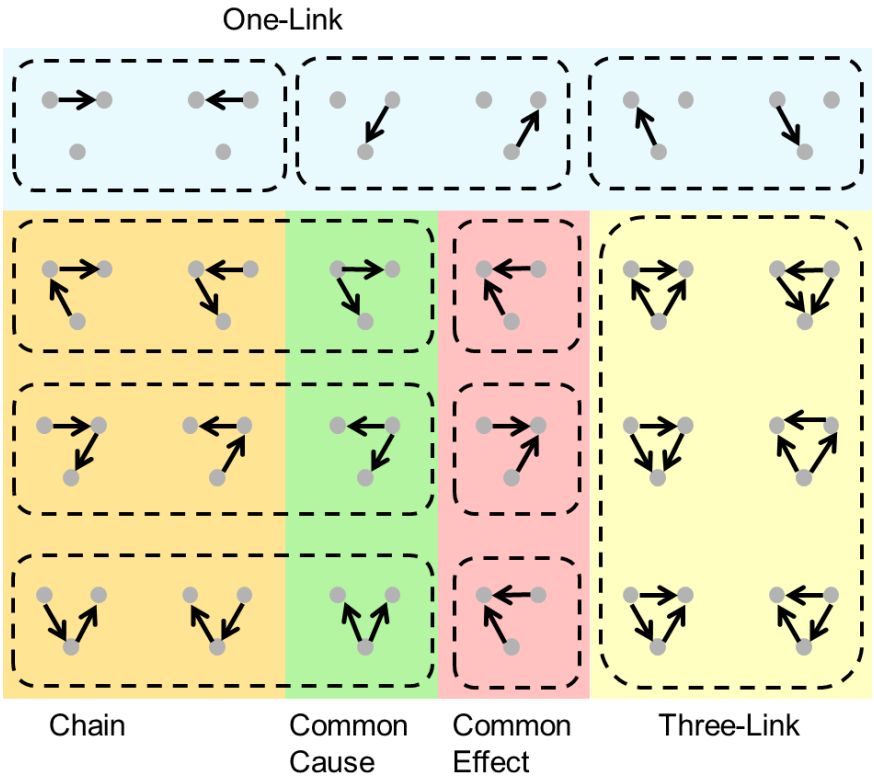

Figure 2. All possible 3 node causal structures. Shaded areas indicate typical descriptive categories, dashed lines indicate Markov equivalence classes.

Imagine trying to convince someone that it is the flagpole that causes its shadow and not vice-versa (Bromberger, 1966). Intuitively, the way to do this is to demonstrate that changing the flagpole – e.g. shaking it, or taking it down – affects its shadow, while changing the shadow - e.g. shining a light to blank it out - does *not* affect the flagpole. This is the essence of active learning about causation. Reaching into a system and holding parts of it fixed, is the basic exploratory move to establish what depends on what.

In terms of the CBN formalisation, these interventions are represented by "link breaking" graph surgery (see Figure 3), whereby the links into an intervened-on variable are temporarily severed and its value is fixed exogenously. People are better at identifying causal structures when they are allowed to intervene on them (e.g. Lagnado & Sloman, 2002, 2004,

2006; Steyvers et al., 2003; Waldmann & Hagmayer, 2005). Further, people are sensitive to the differences in the predictive and counterfactual implications of naturally observed patterns of data versus those resulting from active interventions (Sloman & Lagnado, 2005). This sensitivity is often a necessary part of reasoning and inference (Pearl, 2000), because *imagining* the effects of interventions on our mental causal graphs allows for hypothetical and counterfactual thinking.

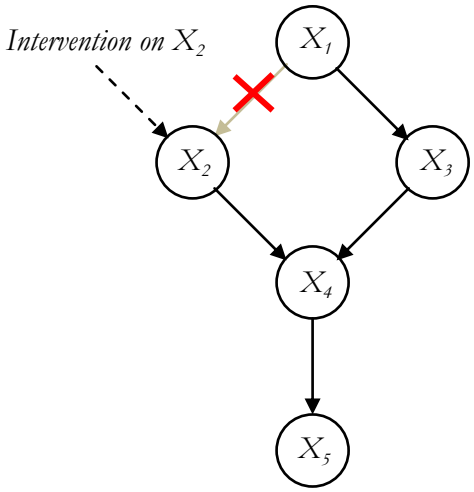

Figure 3. A depiction of an intervention on the DAG from Figure 1. The link from X_1 to X_2 is broken and X_2's value is fixed by the agent from outside the domain of the graph as indicated by the dashed arrow.

Hierarchical representation.

Prima facie, adult causal judgements are influenced by pre-established knowledge about causation. For example: temporal ordering of events (Lagnado & Sloman, 2006); knowledge of the causal roles of objects (Kemp, Goodman & Tenenbaum, 2010); or the implications of their spatial configuration, make certain causal interpretations of situations much more natural than others. What is not so well understood is how people integrate statistical (covariation-based) evidence and non-statistical (visual or temporal cue-based) evidence when making inferences about the causal structure underlying encountered phenomena (Lagnado, 2011; Lagnado & Speekenbrink, 2010). A recent theoretical framework aims to unify our understanding of both types of information. The idea is that human learning and inference may be best understood as a

process of hierarchical Bayesian inference (Goodman, Ullman & Tenenbaum, 2011; Kemp, Goodman & Tenenbaum, 2010; Griffiths & Tenenbaum, 2009; see Clark, in press, for a review), where incoming data is interpreted via layers of generative models of increasing abstraction.

In terms of causal beliefs, the framework suggests we not only have *specific* beliefs about relationships between particular objects; but also: beliefs about the *roles* of types of objects in causal structures; and beliefs about the *types* of structures which underlie broad *domains* or *classes of events* (Figure 4). The hierarchical application of Bayes theorem to each of these levels provides a rational framework, showing how hierarchies of CBNs can be constructed optimally to capture increasingly generalised causal information, though bottom-up induction from observed similarities and patterns across multiple causal learning instances. This high-level causal knowledge, sometimes called *causal grammar* (Kemp, Goodman & Tenenbaum, 2010), then accelerates subsequent inferences at lower levels by biasing the hypothesis spaces of possible or likely structures based on identifying cues to high-level structure. This model of learning and cognitive organisation thus gives a holistic account of why recognised objects and domains can provide top-down cues to the likely causal structures involved in a given environment. The second experiment in this paper will explore the extent to which this framework fits how humans learn about causation when causal grammar cues are available.

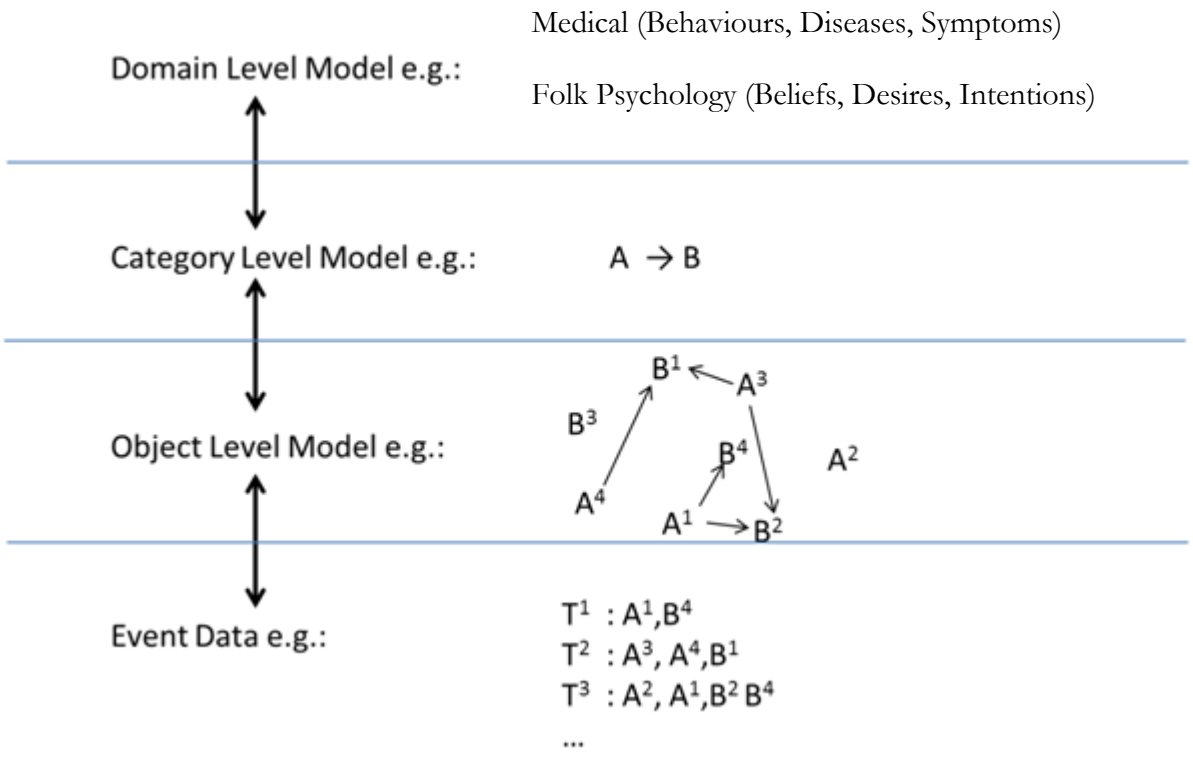

Figure 4. A schematic of proposed hierarchy of our causal knowledge. Parameter estimates and generative probabilities can be included at the different levels but are left out here for simplicity. Note the 'category level' generalisation that As cause Bs. It is clear that such knowledge could *guide* object level inferences about subsequently encountered As and Bs. This is the essence of the framework.

"Optimal" active learning

Active learning is the process of learning about a domain by selectively intervening on it to establish its underlying structure. Different interventions on a system clearly vary in their effectiveness in distinguishing between possible structures. Therefore, in order to assess people's intervention choices and subsequent learning one must find a way to measure the informativeness of different interventions.

We can define an optimal intervention as one which leads to the largest *expected reduction in uncertainty* about the true underlying causal structure given data and what is already known. This makes it possible to calculate optimally efficient interventions for solving simple problems by using Bayes theorem and information theory (Murphy, 2001; Tong & Koller, 2001). These

optima can then be used to assess the performance of different learning strategies and the actions of human learners.

Using information entropy (Shannon, 1951), we can quantify our uncertainty (*U*) about the true causal structure for a problem as

$$U = -\sum_{i=1}^{n} P(H_i) log_2 P(H_i)$$

where $H_{1\ldots n}$ is a hypothesis space consisting of *n* distinct causal structural hypotheses consistent with the number of variables in the current problem. *U* is a measure of how much we do not know about the underlying system. The more peaked a probability distribution across hypotheses, the smaller the value of *U,* and the more certain one can be that the structure with the maximum posterior probability is the actual structure.

In this paper we present a model that uses Bayes theorem to calculate the change in the probability distribution across possible structures, given every possible intervention (I_j) and every subsequent pattern of activation (A_k) of unfixed variables given the variables that are fixed by the intervention. To do this, the *likelihood* of each 'intervention-activation' pattern is calculated, using the known parameters of the underlying CBN then summed over hypotheses to provide the denominator:

$$\sum_{i=1}^{n} P(A_k|I_j, H_i) P(H_i)$$

Finally, the model calculates the posterior probability distribution over hypotheses using Bayes theorem:

$$P(H_i|I_j, A_k) = \frac{P(A_k|I_j, H_i) P(H_i)}{\sum_{i=1}^{n} P(A_k|I_j, H_i) P(H_i)}$$

It is also possible to estimate CBN parameters (link strengths and base rates) contemporaneously with learning structure (Nyberg & Korb, 2006). But, because our focus in this paper is structure learning, we omit this additional step.

Each intervention results in the observation of one of a subset of possible activation-intervention patterns, each with an informational value for how much they would change the model's uncertainty (ΔU) about the true structure and a likelihood for observing exactly those activations given the intervention [$P(A|I)$]. From this, it is possible to calculate the *expected* reduction in uncertainty (EU) of an intervention, using an analogue of the normal expected value formula

$$EU_j = \sum_{k=1}^{n} \Delta U_k P(A_k|I_j),$$

and the definition of uncertainty given above. Thus, this model, when given a probability distribution, returns a vector of expected information gains (one for each intervention).

Strictly speaking, this model is not necessarily *optimal*. It is "greedy" in the sense that it opts to learn as much as possible on each intervention, as opposed to planning several interventions ahead, like a chess player. However, for causal learning, unlike chess, the environment is stationary, meaning consequences of later actions do not depend on the results of earlier actions. For this reason it is not clear that looking several moves ahead would lead to different interventions.

Even without looking ahead, this model is extremely computationally intensive. It must compute, and weigh up, the informational implications of every possible outcome before it picks a single intervention. This becomes computationally intractable for environments with more than about five variables (Scheines, Easterday & Danks, 2007). It should be stressed therefore, that this model is only intended to provide a *computational level* solution to the problems faced by

cognitive agents (Marr, 1982). The implication that we update and reweight our hierarchy of causal beliefs globally based on every minor observation is intuitively implausible, especially given findings that demonstrate the global inconsistency and intransitivity of many of people's beliefs (Gilovich, Griffen & Kahneman, 2002; Kahneman, Slovic, Tversky, 1982). Similarly it seems unlikely that we take all of our existing knowledge into account when selecting an intervention.

Relevant existing studies

Kushnir et al (2009) taught people to draw the causal structure they believed to underlie observed patterns of movements of sticks coming from a box with a hidden mechanism inside. Participants were generally proficient at this task, and would often spontaneously posit the presence of hidden nodes in their causal models when this would better explain their observations. Lagnado and Sloman (2006) also probed how people inferred the structure underlying simple causal systems, on the basis of seeing the outcomes of single interventions on the system in addition to temporal-order cues. They found that people did make use of the implications of these interventions as well as temporal cues in their judgements. Overall, their data suggested a model of causal learning based on hypothesis generation and testing.

Steyvers et al (2003) also probed people's ability to distinguish between competing causal explanations on the basis of a one single-variable intervention of their choice, after a series of observations. They found that accuracy improved following the intervention, although overall performances were fairly poor, possibly because people could only make a single-variable intervention. Steyvers et al also compared participants' intervention choices to that of an optimal uncertainty-reducing test model, and several heuristic models (which assumed participants only considered a restricted hypothesis space when trying to choose a diagnostic intervention). Overall, they found that the optimal model was a poor fit to participants' actions

while the restricted models provided very good fits, again suggesting hypothesis-driven interventions in active learning.

Finally, Kemp, Goodman and Tenenbaum (2010) probed causal grammar learning over multiple learning instances in a *blicket detector* paradigm, by asking participants to sort objects into groups by causal strength, and then use these groups to draw inferences about a novel object. They found that people made appropriate inferences given enough trials and that a Bayesian hierarchical model of learning provided a better fit to participants' inferences than bottom-up or exemplar models.

Research questions

This paper will investigate causal learning mechanisms, looking at how people learn more complex causal models (with up to 6 variables); when they are allowed to make multiple interventions; and to intervene on more than one variable at once (unlike previous research which focuses on one-variable interventions); and when causal structure is also present at the level of a causal grammar over multiple problems. We expect that for the more complex problems people will need to use heuristic strategies to cut down the search space and will use their ability to fix variables off to break down the problems into manageable parts. Potentially, we develop sophisticated belief structures by building up and amending our beliefs *locally*, focusing on one or few potential causal links at a time, or making inferences at only one hierarchical level rather than across all of them (Fernbach & Sloman, 2009). Simple but robust active mechanisms may be sufficient for selecting interventions to perform which, in the limit, allow us to construct reasonably accurate causal representations of our environment. Implicit sensitivity to *basic* statistical information (Lewicki, Hill & Czyzewska 1992), probably feeds this process but, arguably, some more explicit, focused organisational mechanism is also required.

The experiments in this paper are intended to examine what people actually do when learning about causal structure so as to get a clearer picture of the mechanisms involved. In particular, we investigate the basic learning mechanisms that people use when only interventional or observational data is available, and to what extent the presence of a simple cue across instances impacts on this learning behaviour. The guiding intuition is that an understanding of causal learning at its most stripped down, and hierarchical inference in its most basic form, will provide a clearer picture of how causal learning in humans works in complex and applied domains.

General Method

Overview

In order to investigate the above questions, we designed a causal learning task, based loosely on the 'recover the graph' paradigm used by Lagnado and Sloman (2006), Kushnir et al (2009) and the rational framework of Steyvers et al (2003). The paradigm involves asking people to identify the structure of a system, by drawing its directed graph, based on observations and interventions.

Extensions to previous work

These experiments represent an extension to previous 'recover the graph' studies in several respects. Kushnir et al. (2009) and Lagnado and Sloman (2006) looked only at learning of 3 and 4 node causal structures, and in the former case only by way of observational information and their learning problems were framed within realistic causal scenarios. This means that established knowledge was likely to play a large part in inferences. Participants were asked to complete the causal structures by drawing links between these objects on a paper diagram in the case of Kushnir et al, and by answering multiple specific questions in the case of Lagnado and Sloman.

In the current experiments, problems ranged from simple to complex with up to 6 nodes and 5 causal links, to examine how learning scales up. Causal learning problems were also deliberately posed without the provision of cover scenarios to avoid, as far as possible, the influence of established causal knowledge. Participants were able to perform interventions, and the endorsement of causal links between nodes takes place in the interactive computer environment during the task, rather than afterwards, to minimise memory effects and provide more fine-grained information about learning.

Steyvers et al (2003) used an optimal learner model to investigate a single simple intervention choice made after a series of forced observations. Our experiments extend this by allowing participants a free rein to perform simple and complex interventions; to do so over as many tests as desired, and to do so without a fixed observational period at the start. This makes it possible for us to investigate what patterns of contemporaneous and sequential interventions people use to uncover a causal structure, how much they try to achieve on each test and to what degree they break the problems down into simpler parts. It also allows us to build more sophisticated models of active learning, taking account of the cumulative results of multiple interventions.

Instruments and materials

The experiments were programmed in MATLAB Version R2010a, and the interface utilised MATLAB's Psychophysics Toolbox Version 3. Full instructions are in Appendix A. The experiments were followed by a paper-based questionnaire (Appendix B).

Procedure

The participants' task was to establish and endorse the causal structure underlying patterns of activation of a group of 'nodes' which appeared as filled coloured circles on the

screen. Participants completed multiple problems, and in each problem they were allowed to perform as many tests as they wished. On each test, the program determined which 'nodes' would activate, probabilistically, according to a hidden causal Bayes net with a structure corresponding to the block and randomised problem number. The parameters of these causal Bayes nets were always the same: each node activated at random with a base rate of 0.1 unless it was "fixed" by the participant (see below). Where there was a causal link in the underlying structure, an active cause node would have a 0.9 probability of activating its immediate children. Participants were informed of these parameters as percentages in the instructions. Causes were always generative (i.e. they always raised the probability of their effects). Multiple causes were integrated by the generic noisy-OR equation. The underlying structures were always acyclic and each test was independent of previous tests.

Performing tests. In each problem, participants were initially presented with a white screen with between 3 and 6 randomly positioned grey filled circles representing the nodes (Figure 5). In each new problem, new circles would be drawn in different random positions.

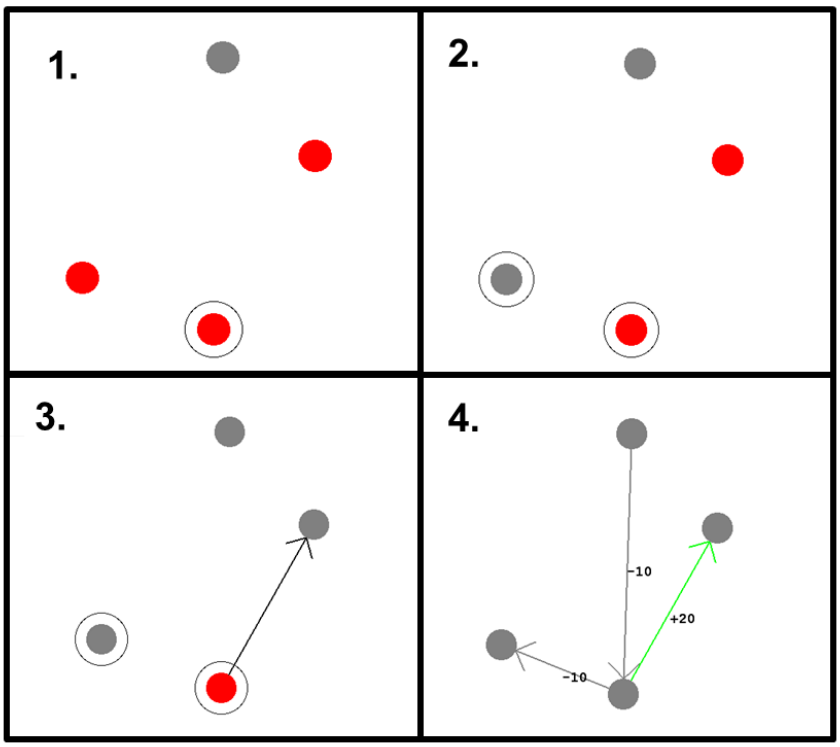

Figure 5. User interface. 1. 4 node problem, bottom node is fixed "on", test is in progress and two nodes middle nodes have activated but top node has not. 2. Another test in which left node is switched off. This time only right node activates. 3. Participant has now endorsed a link from bottom to right hand node. 4. Feedback screen in shich participant sees which links were right, wrong and missed, and receives a score for the problem.

Participants tested the network with the spacebar. Each press constituted one *test* where "activated" nodes would turn red for 1.5 seconds while the others would remain grey. At any time, participants were able to *fix* nodes by clicking on them. Each node could be fixed either to be *on* or *off* on subsequent tests. This was indicated by the addition of a black circle around fixed nodes, with *on* nodes staying permanently red (Figure 5). Initially, all the nodes were *unfixed*. Nodes did not reset to *unfixed* after each test but stayed in their selected state until they were clicked on again.

Fixed nodes behaved as ideal link-breaking interventions, temporarily altering the underlying Bayes net by severing any incoming links and fixing that node's *activation* to "on" or

"off". Participants could continue intervening and testing until they indicated that they were finished with a problem.

Endorsing structures. Participants chose the causal structure for each problem by marking links as they went along, clicking between nodes to endorse a causal link. In this way, participants could effectively "draw" the structure that they judged responsible for the data, link by link. One click between two nodes made an arrow appear (in a randomised direction), a second click reversed its direction, and a third removed it again.

Receiving feedback. After each problem there was a feedback screen that displayed the true causal structure overlaid with the participant's endorsed arrows (Figure 5) and a breakdown of their performance for that problem. Payment was tied to performance, and the scoring system was designed to create a realistic incentive structure for learning. Participants received: +20 points for each correctly endorsed arrow; -10 points for each wrongly endorsed arrow; -10 points for each missed arrow; -1 point for each test performed.

Questionnaires. At the end of block 4, participants were thanked for participating and the program ended. They were then given a short questionnaire to explain how they had solved the task and to describe any specific strategies they had used (including a diagram if possible).

The "optimal" active learner model

Each problem in Experiment 1 was modelled using the "optimal" Bayesian learning model, starting with a uniform prior across possible hypotheses. A minimum description length prior (Hansen & Yu, 2001) was also considered, to reflect the common assumption (e.g. Sprites, Glymour & Scheines, 1992/2000) that we are biased to endorse sparse but powerful causal

structures[1]. Ultimately though, it was decided that participants in a causal learning experiment would expect a variety of link numbers and structural patterns, making a uniform prior the natural choice.

There were: 24, 542 and 3.78 million permissible directed acyclic graphs with 3, 4 and 6 nodes respectively, making calculations for the first blocks trivial relative to those for the second half. There were 19, 65 and 665 possible intervention patterns in blocks 1, 2 and 3/4 respectively, each with 2^n possible activation patterns where *n* is the number of unfixed nodes in that intervention As an example, an intervention which fixes one node to "active" in block 1, leaves the other two nodes (A & B) unfixed, and so makes a set of four possible 'activation-intervention' patterns: 1. *neither active, 2. A active, B inactive, 3. A inactive, B active, 4. both active*. Each of these has a distinct likelihood under each hypothesis. Interventions in which all of the nodes were fixed were not included because they always had an informational value of zero. The structure in which none of the nodes were causally connected was also not included in the hypothesis space because the task instructions implied that there would be at least one causal connection in each problem/structure. The super-exponential increase in the calculations required to select interventions meant that it was not feasible to fully model blocks 3 and 4[2]. Formal analysis using the benchmark model is therefore mainly restricted to blocks 1 and 2.

To establish the 'informativeness' of different interventions for learning about the structure of a causal system, the benchmark learner model was tested 1000 times on the three

---

[1] Such a bias makes intuitive sense when you consider that structuring data causally is all about simplifying inferences. There would be no such advantage if we learned to learn structures in which every variable is linked to every other one.

[2] After one week of processing on a high-end PC (8-core, 12 Gb RAM, Linux Desktop) the model had only just selected its *first* intervention for problem 1 in block 3.

and four node problems. This allowed us to measure the frequency with which different interventions were selected, as well as the typical patterns of interventions and corresponding shifts in the probability distribution. In these tests, the model was parameterised to keep performing tests until the posterior probability of the most likely structure was greater than 0.9 and then to endorse that structure. Uncertainty typically decreased sharply in the last few tests so the model is robust to small changes in this value. The results of these tests are summarised in Table 1.

What Interventions did the Model Select? For three-node learning, *fixing one active node* at a time was the most frequent choice of the benchmark learner model at (70.58%; Table 1). The only other type of intervention frequently chosen was one with *one node fixed on and another fixed off* (28.77%). For four node problems, the modal intervention type was to fix *one node on and one off* (45.58%, Table 2) followed by fixing just *one on* (36.89%). However, also selected were: '*two-on, one-off*' (12.35%), *'one-on two-off'* (4.51%) and '*two-on'* (0.67%). While it was not feasible to do this analysis for six node problems, the expected values (Table 3) that were computed for the first intervention gives an idea. Fixing '*one-on'* was the most informative move from a flat prior but '*one-on, one-off*', '*one-on, two-off'* and *'two-on, one-off'* followed close behind suggesting that they were also likely to come into play as the probability distribution evolved.

Possible intervention types for structures with three nodes	No. possible outcomes	EU for first test (bits)	Frequency chosen By Bayesian learner over all tests (%)
All three nodes free	8	0.2016	0
One node fixed on, two nodes free	4	1.0165*	70.58*
One node switched off, two nodes free	4	0.0677	0
One node on, one node off, one free	2	0.4872*	28.77*
Two nodes off, one free	2	0.0025	0

Table 1. The expected value column shows how informative these interventions are for updating from a flat prior (so on the first test for each problem). As the probability distribution becomes more peaked, these values shift around, which is why the model does not always select the same intervention.

Contrary to the claim (e.g. Mill, 1848; Steyvers et al, 2003), that starting with passive observations is a rational way to initially narrow down the hypothesis space of possible structures, in this task, the 'nothing fixed' intervention was *never* selected by the model. This was due to the high probability of no nodes activating by chance (.73 block 1, .65 in block 2 and .53 in blocks 3 and 4). The chance of no nodes activating, if everything is unfixed, is given by the product of the probability that each node fails to activate at random, (e.g. 0.9^n where n is the number of nodes). When no nodes activate, nothing can be learned about the underlying structure, making passive observation a relatively uninformative move, at least when intervening carries no additional cost.

Possible intervention types for structures with three nodes	No. possible outcomes	EU for first test (bits)	Frequency chosen By Bayesian learner over all tests (%)
One node on, three free	8	1.3094*	36.89%*
One node on, one off, two free	4	0.7749	45.58%*
Two nodes on, two free	4	0.9550	0.67%
Two nodes on, one off, one free	2	0.3825	12.35%*
One node on, two off, one free	2	0.2741	4.51%

Table 2. The only interventions selected by the model for four node problems. The other five, generally less informative, interventions are excluded.to save space.

When were the different interventions used and what role did they play? Figure 7 shows a typical pattern of the model's intervention choices (and subsequent observations) to illustrate how these interventions allowed the model to home in on the correct structure efficiently. Problem 3 in block 1 was selected as exemplar of this inferential process.

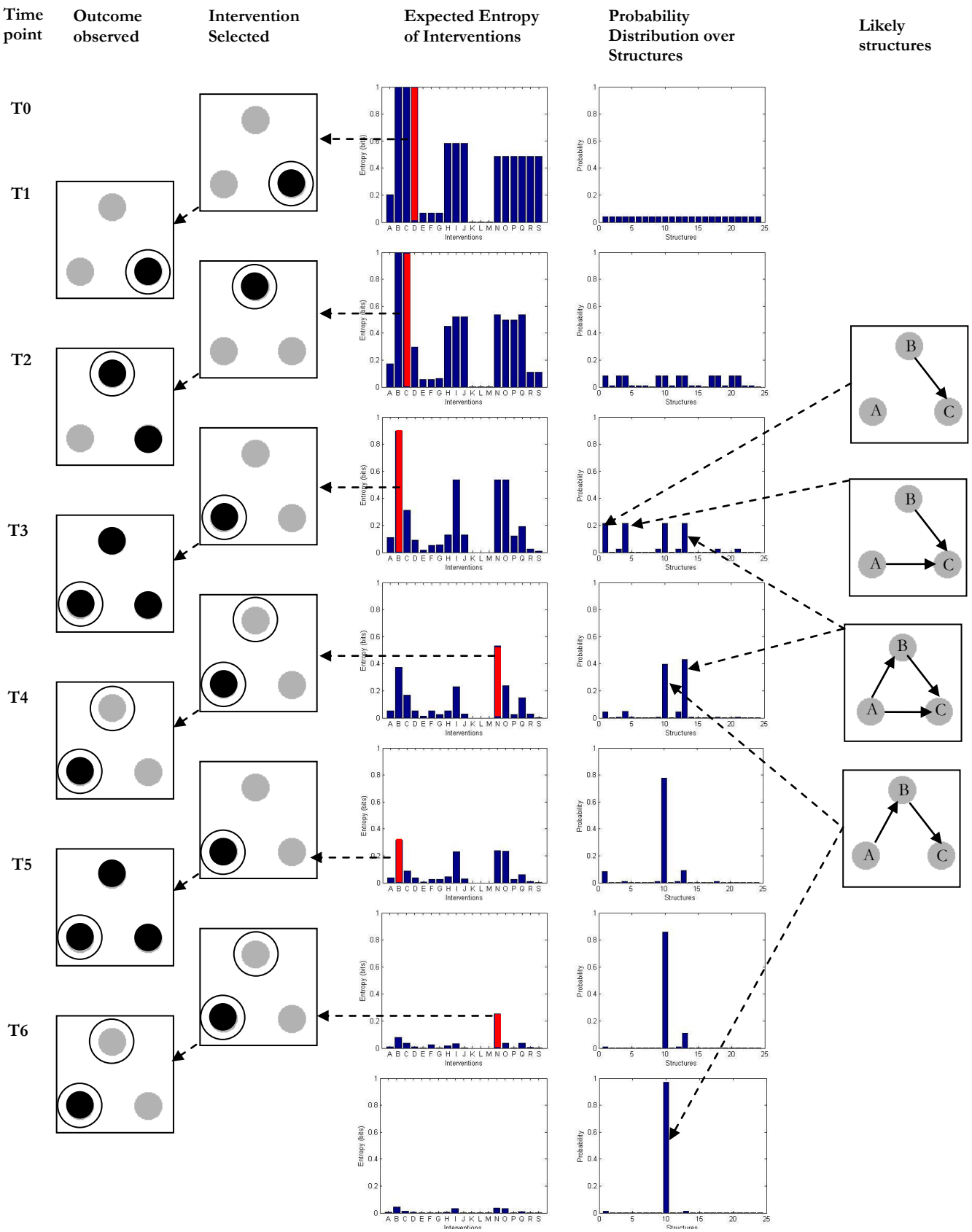

Figure 6. Schematic of Bayesian learner model identifying the chain structure in problem 3 of block 1. Initially (T1-T3) single positive interventions are the most informative but after T3 a disambiguating double intervention is the most informative as it distinguishes between the triply connected structure and the chain structure. The model stops and endorses the chain structure when it reaches .9 probability.

On average the model took 6.4 tests to reach 90% certainty about the three node networks, and 11.0 for four node networks. The model's choices, at least in block 1, can be explained in terms of the discovery of a few hypotheses, and then subsequent discrimination between these hypotheses. Thus, the two common intervention patterns from Table 1 serve distinct roles. The model always starts by performing tests with each node individually fixed to *on*, one at a time. Each corresponding activation leads to a big rise in the probability of structures that contain a link from the fixed node to those that activate. This narrows the model's search down to the few structures consistent with all intervention-activations so far. Outcomes where two free nodes activate are less discriminating. This is because two activations could be due to a structure with: two direct links (i.e. the fixed node is a "common cause"), a "chain" (i.e. where one of the free nodes is a direct effect and the other is an indirect one), or one with two direct links *and* an indirect one. Thus, in cases where multiple activations occur, the most informative interventions become those that effectively disambiguate between these options by testing for the presence of a direct link between the current fixed node and one of the free ones. This is achieved by complex interventions that fix all but one of the free nodes to be *off*, thus testing for a direct connection between the *on* node and the remaining unfixed node.

Something akin to this pattern continues when it comes to four node problems. Fixing single nodes to *on* remains the most informative intervention initially. Similarly, double and triple activations necessitate follow-up interventions, which fix all but one of the resulting activated nodes to *off*, in combination with the same fixed *on* node. However, the hypothesis space is exponentially larger for four nodes making it difficult to represent some of the more complex effects of observations on the probability distribution. Certain interventions, which the model occasionally selects, e.g. where more than one node is fixed "on", change the probability distribution in a way that is too complex to lend itself to a simple intuitive explanation.

Participants.

Twenty-four members of the UCL psychology subject pool took part in Experiment 1 (13 male; 18–50 years age range), and another 24 took part in Experiment 2 (12 male; 19 to 47 years age range). They received between £4 and £6 based on their performance. Nine participants were excluded from the final analysis (five in Experiment 1 and four in Experiment 2) because it was clear from their behaviour in the task and questionnaire responses that they had not fully understood the instructions.

Experiment 1

Method

Design. In the first experiment we tested causal learning without any additional cues. Participants all saw four blocks of five problems making a total of twenty causal structures to be learned. The whole task took around 45 minutes to complete. Problems were randomised within blocks, but the order of the blocks was the same for all participants. Blocks 1 and 2 were made up of 3-node and 4-node problems respectively while blocks 3 and 4 were made up of six node problems.

Stimuli. The causal structures tested are shown in Appendix C. In the experiment the position of the nodes were randomised so that they would appear in different positions on the screen for different participants.

Results and Discussion

Overall performance.

Performance in Experiment 1 was highly varied. Final scores for the 19 participants included in the analyses ranged between +661 and -1021 with a mean of 93.42 (SD= 505.41). Participants were split into two groups based on their performance in the first two blocks of the

experiment using a *kmeans* clustering procedure (MacQueen, 1967) in order to facilitate analyses. Eleven participants were classed as high performers, scoring 256.47 (SD= 278.31) points and finding 49.18 of the total 66 links (SD=9.23) and wrongly endorsing a further 30.62 (SD=17.74). The other eight were classed as low performers, with a mean overall score of= -154 (SD=300.841), finding 30.50/66 (SD=10.28) causal links but wrongly endorsing a further 57.62 (SD=26.77). The low performers did not, on average, perform significantly better than chance (p=.44). Chance performance can be interpreted variously for this task. Neither performing tests nor adding links would leave a participant with -10**Missed Links*=-660 points, while randomly endorsing structures without performing any tests, scores -909. By performing tests and still adding links at chance, it would be easy to do significantly worse than this. However, the first, most conservative, measure of chance was used in order to exclude the possibility of artifactual above-chance scores.

Performance in different problems.

There were significant differences in the number of tests performed per problem as the number of variables involved increased. For the three variable problems in block 1, participants performed 6.14 (SD= 1.23) tests on average before endorsing a structure, with 10.23 (SD=0.63) for the four variable problems in block 2 and 17.05 (SD=1.84) for the six variable problems. An omnibus test for between-block differences in tests performed was significant, $F(1,18)=64.55$ $p<.001$, and post hoc tests revealed significant differences between all blocks (apart from 3 and 4 which had the same number of variables in each). There were no significant differences between number of tests required for the different problems within blocks. There were also no significant differences between the number of tests performed by problem for high performers compared to low performers, nor was there an interaction between number of tests performed by problem and whether participants were high or low performers.

Scores went down slightly as the problems got harder, $F(1,19)=4.89$ $p=0.013$, with blocks 1 and 2 both significantly better than block 3 ($p=.013$ and $p=.05$ respectively), but it is not possible to draw strong conclusions from this, as the score variance increased dramatically with the number of variables in a problem (SD=24.29 for block 1 and 48.25 for blocks 3 and 4). The only significant difference between scores for different problems with the same number of variables were in block 1 [$F(4,15)=5.42$ $p=.001$] and this appeared to be due to difference between the one link problem and the three link problem, means -12.1 and 25.26 respectively, $p<.001$. This difference is to be expected because in problem 1 scores were bounded at 20 points (one correct link) whereas in problem 5 they were bounded at 60 points (three correct links). Time taken per problem did not vary significantly across the blocks or problems, nor did it vary between low and high performers or correlate with overall score.

Efficiency and sensitivity

By looking at participants' actions at each time point, in relation to the evidence they had seen by that point, it was possible to measure how efficient they had been at performing interventions. This was achieved using the optimal active learning model, by computing the posterior probability distribution over structures given the interventions and observations seen by each participant at each time point, and then calculating the expected reduction in entropy of the different possible interventions for their next step. This could then be compared to the intervention chosen by the participant. From this it was possible to say that high performers selected the most efficient intervention on 23.9% (SD=15.7) of tests in blocks 1 and 2, averaging 53.5% (SD=11.4%) of the possible entropy-per-test. Low performers only selected the most informative intervention 9.7% of the time (SD=9.3%) but still averaged 43.6% of the possible entropy per test. Selecting actions at random one would only select the most efficient

intervention 3.4% of the time, putting the high performers, but not low performers, significantly above chance in this regard as well (p=.002 and p=.11 respectively).

It was also possible to look at how effectively participants had distinguished between structures by the time they decided they had finished a problem. This can be assessed by looking at how biased participants' final posterior probability distributions were. High performing participants endorsed structures when what they had seen justified having a subjective degree of belief of just .30 (SD= .10) in the most likely hypothesis being correct while low performing participants were even lower at just .20 (SD=.099). This suggests that, in general, participants were endorsing structures prematurely, before they had enough information to have a high chance of being completely correct.

Strategy insight

In order to try to establish what was behind the large individual differences in performances between participants, the questionnaire at the end of the experiment probed whether participants had, or thought they had, insight into how they gone about identifying causal links.

11 of the 19 participants in Experiment 1 were able to describe how they thought they had solved the task and these descriptions were later categorised in two general strategies:

Simple link endorsement (n=5)

"Fixing each node at a time, and finding its effects" (Participant 7)

"Designated [..] nodes as fixed active and tried to observe if they caused other nodes to respond"(8)

"activated each node once to see what nodes it affected and assumed links" (18)

This was formalised as the following procedure: 1. Fix one node on and test it one or several times. 2. Add a causal link from the fixed node to all other nodes that activated on the majority of tests. 3. Repeat with each node (Figure 7).

Simple link endorsement plus additional disambiguation step (n=6)

The first step was the same as simple link endorsement with the additional step as described below:

"..Inactivating nodes during this trial to establish that nothing else could be activating the other node other than the one under test" (12)

"I would fix some nodes as inactive to test if the links were only present just because some are active"(15)

"..placed a node on and viewed any results. If multiple nodes were turned on as a result, the paths were determined by blocking one of the activated nodes (turning it off) and viewing whether the same results was obtained."(23)

This was formalised as an additional step which was performed whenever more than one unfixed node activated: 4. Perform one or several tests with the same fixed 'active' node *and all but one of the previously activated nodes fixed to 'inactive'*. If the unfixed node does not activate, remove the causal link added in step 1. Repeat this procedure with each of the activated nodes (Figure 7).

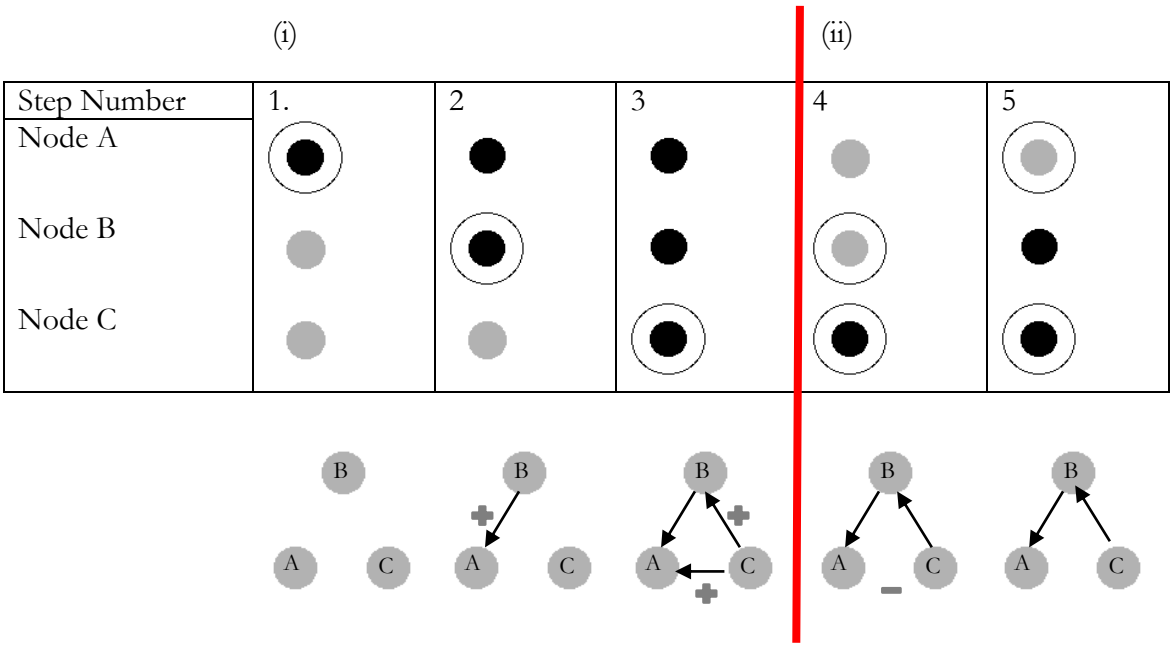

Figure 7. (i) Illustrates 'simple link endorsement' on a three variable problem, each step can be repeated multiple times (see below). (ii) further illustrates a disambiguation step, resulting from the double activation in test three. Below are the causal links endorsed by these strategies as a result of each step. For the 'simple link endorsement' strategy, the model would stop at the red bar while for the 'disambiguation' strategy it would also perform the additional step and therefore remove the link from C→A resulting in a chain structure.

Fitting active learning models

While many participants' appeared to have insight into their own active learning, it was also necessary to establish to what extent these descriptions reflect what participants' actually did in the task. For this reason we performed a model comparison, fitting and comparing the two models, as well as two baseline models and a Bayesian model, to participants' actions and causal link endorsements during the task.

At each time point, in each problem, each participant would choose what intervention (or *action*) to perform and had the option to update the causal links they had marked in (thus providing a proxy for their *belief state* about the true causal structure at that time point). This meant that the interventional action taken, combined with the causal links currently marked, provided a (multivariate) dependent measure for the model fit.

We used maximum likelihood (Myung, Forster & Browne, 2000) to fit the models to the data from each individual participant (for full specification of the models and full results see Appendix D). The models all had between 0 and 3 free parameters. We then computed the Akaike Information Criterion (AIC; Akaike, 1974) and Bayesian Information Criterion (BIC; Swartz, 1978) for each model, for each participant in order to compare their fits. The parameters used in the models were:

The models tested were as follows:

1. **Random:** Assumes that people selected every intervention at random as well as selecting each belief state at random. Parameters - None

2. **Simple endorsement**: As described in the previous section. Assumes that people always act on single nodes and always update their *belief state* to include any resultant activations. Parameters -

W_A - Probability of *acting* in accordance with the model at each time point.

W_B - Probability of holding a *belief state* in accordance with the model at each time point.

W_R - Probability of repeating the previous *action* or the previous *belief state*. This was included because many participants reported performing each test several times in the questionnaire: 2 times (n=3), 5 times (n=2), several times(n=1) and the mean number of tests performed of 251, implies 2 to 3 tests per step.

3. **Disambiguation:** As described in the previous section. Assumed that people would behave the same as in simple endorsement, except after observing $n>1$ activations. Where, on the following trial(s) it assumes people perform disambiguation steps (keeping the same node fixed on and between 1 and n-1 of the activated nodes fixed off) and remove links when they are not corroborated. Parameters – W_A, W_B, W_R.

4. **Random with repetition:** Assumed that people select actions and belief states at random but that they were also likely to repeat an action or fail to update a belief state. Parameters – $W_{R..}$

5. **Bayesian:** Assumed that people would select actions with a probability dependent on the expected reduction in uncertainty of that action, and select belief states with a probability dependent on the probability of that structure at that time point according to a Bayesian learner exposed to the same data. Because it was not possible to compute the expected reductions in uncertainty of the interventions in blocks 3 and 4, this model could only be run on blocks 1 and 2.

W_a – "Temperature" for "softmax" function converting an action's expected value to the probability that it would be selected.

W_β – Temperature for softmax function converting the posterior probability of a structure's being correct, to the probability that it would be selected as the belief state at that time point.

W_n – A noise parameter capturing the probability of selecting an *action* with zero value, or a *belief state* with zero probability (see Appendix D for full justification of the use of these parameters).

The results of these model fits to blocks 1 and 2 are summarised in Tables 3 and 4. According to both the AIC and the BIC, the *disambiguation* model provided was the best overall fit. It had the lowest BIC score for all 11 high performers and 4 out of 8 low performers. The remaining low performers were better described by the *random with repetition* model. Additionally the *simple endorsement* model was a better fit for all but one participant than the *Bayesian* model which proved to be a fairly poor fit to most participants. High performers also tended to have stronger parameters for acting and updating beliefs in accordance with the *simple endorsement* and

disambiguation models than did low performers. Although it was not possible to test the *Bayesian* model on blocks 3 and 4, it was possible to test the other four models. These tests showed that the trends continued with 16/19 participants still best described by the *disambiguation model* looking at all four blocks (see Appendix D). Once again, all of the high performing participants are best described by the *disambiguation* model and just three of the low performers are again best described by the *random with repetition* model.

Model	Maximum Log Likelihood	sd	Akaike Information Criterion (AIC)	sd	Bayesian Information Criterion (BIC)	sd	# participants best described
Random	-1270.11	643.68	2540.22	1287.36	2540.22	1287.37	0
Simple	-528.06	256.36	1062.12	512.72	1072.33	513.94	0
Disambiguating	-402.18	245.80	810.35	491.59	820.56	492.64	15
Random (repeats)	-593.14	260.72	1188.28	521.44	1191.68	521.72	4
Bayesian	-773.32	373.15	1541.44	744.25	1490.18	717.74	0

Table 3. Measures of fit for the different strategies, averaged over participants

Model	High/Low performers	W_A (W_α)	sd	W_B (W_β)	sd	W_R (W_n)	sd
Random	High	x	x	x	x	x	x
	Low	x	x	x	x	x	x
Simple	High	0.64	0.16	0.80	0.10	0.52	0.19
	Low	0.33	0.25	0.75	0.16	0.70	0.18
Disambiguating	High	0.76	0.10	0.92	0.02	0.65	0.13
	Low	0.33	0.25	0.80	0.14	0.77	0.12
Random (repeats)	High	x	x	x	x	0.46	0.17
	Low	x	x	x	x	0.61	0.22
Bayesian	High	0.23	0.41	0.08	0.19	0.24	0.09
	Low	0.02	0.03	0.03	0.07	0.19	0.11

Table 4. Mean parameter estimates for low and high performers for each of the models

Comparing simple endorsement and disambiguation

While there was a trend, there were no significant differences in performance between those who did not report using a strategy and the purported simple endorsers and disambiguators [mean scores: *no-strategy* -351.8, *simple endorsement* 28.0, *disambiguation* 174.4 $F(2,17)=2.171$ $p=.14$].

A pure simple endorsement model and a disambiguating model were tested independently on problems in the experiment (100 times each, with 1:5 repetitions) for comparison. The results are summarised in Tables 5 and 6. In these tests, the simpler strategy actually performed marginally *better*. This was due to its savings in efficiency. It took 285 tests to achieving 697.2 points on average (repeating each step three times) compared to 476 tests to achieve 657.7 for the model with the disambiguation step added (see Tables 6 and 7). The simple link endorsement strategy was also more efficient on average, and picked the optimal intervention more frequently, but endorsed structures while the maximum posterior probability was lower, and selected the most probable structure given the data less frequently (Table 8). These results are reflected in participant's performances with simple endorsers making more false endorsements (48.17 vs. 30.8) but selecting more efficient interventions on average (23.8% of available entropy vs. 15.2%) and selecting the optimal intervention more frequently (60.0% of the time vs. 48.1%).

This suggests that the advantage of performing the second 'disambiguation' step was that it reduced the chance that extra over-determining causal links would be endorsed by mistake. It did this by effectively pruning out most of the links that had originally been endorsed due to their *indirect* causal influence via other nodes. If the payoffs had been weighted more towards accuracy than frugality, the more complex strategy would have paid out higher. For example, if

the cost of wrongly endorsing an arrow was as high as the payoff for correctly identifying one, *disambiguation* would have scored 539.06 to *simple endorsement*'s 400.50.

No of times each test performed:	Mean Score	Splits by Block:				Tests Req:	Causal Links Right,	Wrong,	Missed
1	514.4	129.40	143.60	120.20	121.20	95.00	60.44	54.38	5.56
3	697.2*	131.80	158.60	200.40	206.40	285.00	64.63	29.67	1.37
5	623.2	110.60	127.90	189.10	195.60	475.00	65.65	21.13	0.35

Table 5. Performance of the 'simple link endorsement' strategy. Performing 2 or 4 tests were also tested for the disambiguation strategy but the result was that, in cases where 1/2 or 2/4 tests resulted in activations, a decision had to be made whether to round up or down. Both options were tested, but performance was significantly worse in both conditions than those with an odd number of tests per node. Therefore, only the 1, 3 and 5 test conditions are included here.

No of times each test performed:	Mean Score	Splits by Block:				Tests Req:	Causal Links Right,	Wrong,	Missed
1	642.0	129.45	166.91	172.04	173.59	180.82	56.79	22.08	9.21
3	657.7 **	122.06	153.65	187.55	194.43	475.87	63.74	11.86	2.26
5	468.1	89.86	88.63	144.43	145.17	753.58	65.41	8.07	0.59

Table 6. Performance of the 'disambiguation' strategy.

Interim summary

In summary, Experiment 1 demonstrated that many people are able to solve complex abstract causal learning problems (although many others struggle). Behaviour of successful participants in this experiment was better explained as an iterated process of local testing and disambiguating than one of trying to maximise entropy over all possible hypotheses.

Experiment 2

The experiments of Kemp, Goodman and Tenenbaum (2010) suggested that people can identify and utilise causal grammar information over multiple instances to guide inferences. However, they did not investigate whether domain level information might affect *active* learning,

or how attention might be allocated between learning at different hierarchical levels. Therefore Experiment 2 provides a version of the task in Experiment 1 that allows investigation of these questions.

Method

Design. The first two blocks of Experiment 2 were identical to those of Experiment 1, providing an opportunity to replicate the findings from Experiment 1. Blocks 3 and 4 meanwhile, were similar to blocks 3 and 4 in Experiment 1 but with one addition. Instead of the nodes being grey, they were one of two colours (see Figures 8 and 9 and Appendix C). The questionnaire also contained an additional two questions pertaining to the colours. Participants were asked if they had found the colours helpful in blocks 3 and 4. If they answered 'yes', they were asked to describe the roles of the yellow and blue nodes in block three, then green and orange nodes in block four.

Stimuli. In block 3 there were always 3 blue and 3 yellow nodes, and in block 4 there were always 3 orange and 3 green nodes. Participants were told after block 2 that their task would be the same as before but that: "[they] may find the colours helpful and may be asked about them afterwards."

Figures 8. The two colours of node in blocks three and four. Solid arrows indicate the types of causal link that were present while the dashed arrows indicate those which were absent.

The colours actually corresponded to the underlying structures so as to provide a helpful cue (see Figure 3). In block 3, causal links always originated from blue nodes and acted on yellow nodes; there were no causal links between blue nodes, between yellow nodes or from yellow nodes to blue nodes. In block 4, causal links always ran between nodes of the same

colour, so there were no causal links from green to orange or from orange to green. However, for these cues to guide causal inferences later in block 3 or 4 they first had to be identified.

Results and Discussion

Overall Performance

Final scores for the 20 participants in Experiment 2 ranged between +845 near-optimal and -887 (mean 11.4, SD 576.6), which was not significantly different from Experiment 1; $F(1,39)=0.37$, $p=.55$. They also performed about the same number of tests in total over the experiment as those in Experiment 1 (267.05 / 251.84).

Ten participants were classed as high performers (based on the first two blocks and the procedure used in Experiment 1), overall scoring = 519.5 (SE= 176.55) points and finding 57.17/66 links (SE=5.48) and wrongly endorsing a further 15.9 (SD=8.67). The other ten were classed as low performers, with a mean overall score of= -496.7 (SD=311.56), finding 27.1/66 (SD=11.04) causal links but wrongly endorsing a further 48.6 (SD=28.47). This low performing group were, again, not on average significantly better than chance ($p=0.13$).

Colour cues

High performers, measured based on blocks 1 and 2, did do better in blocks 3 and 4 in Experiment 2 than in Experiment 1; with block 3 and 4 means of 71.0 (SD=238.2) and 285.2 (SD=150.3), $F(1,20)=5.90$, $p=.025$ for the high performers in Experiment 1 and 2 respectively. This provides evidence that the colour cues played some supportive role.

Questionnaire responses were coded according to the ability of participants to describe the role of the colour cues in blocks 3 and 4 after completing the task. Based on this coding scheme, colour cues appeared to help participants to solve the problems in blocks 3 and 4, *if* they were able to recognise their value. Ten participants were able to describe one or both colour cues. Of these participants, the three who were able to describe both, averaged 317.7 (SD 165.5)

points in blocks 3 and 4, the four who could only describe the block 3 cue averaged 57.0 (SD 335.4), the three that noticed only the block 4 cue averaged 133.0 (SD 61.0). The other 10 participants who could describe neither averaged -261.4 (SD 432.1). Thus, there was a significant positive effect of noticing either or both colour cues compared to not ($F(1,19)=7.198$ $p=0.015$) in Experiment 2 as well as a beneficial between-Experiments effect for high performers.

Of the ~3.8 million possible six node causal structures, only 570 were made up of only links from any three nodes to the other three, and only 511 were made up of only links *within* exclusive triples (see Figure 16). This meant that identification of the role of the colour cues in this task reduced the number of possibilities in blocks 3 and 4 by a factor of around 7000.

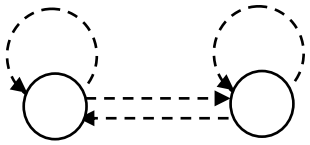

Figure 9. The dotted arrows indicate different possible causal links between two *classes* of object. Reflexive arrows indicate causation within members of the class.

While it was not possible to compute optimal active learning for the full hypothesis space, it was possible to compute with the restricted hypotheses spaces given the grammar. For blocks 3 and 4, the benchmark model averaged 9.6 and 9 trials respectively to endorse each structure, and the expected values of different interventions were markedly different from those computed from the full hypothesis space (see Table 7). Essentially, the model always fixed all three blue nodes in block 3, usually 'two-on, one-off' (56.25%) or 'one-on, two-off' (39.58%). Interventions which fixed yellow 'known effect' nodes were always worse than those that were the same but with the yellow nodes unfixed. Essentially, fixing a known effect *always* reduced the amount you could learn on that trial. In block 4 the model's behaviour was best understood as

treating the problem as two simultaneous three-node problems, performing the same patterns of interventions as in block one simultaneously, *within* each colour.

Intervention type	Outcomes	EU full H space flat prior.	EU if cue known Block3	Freq. Selected by model	EU if cues known Block 4	Freq. selected by model
All Free	64	1.014	.435	0	0.398	0
One On	32	2.660**	1.589 / .289	0	1.215	0
One On, One Off	16	2.079*	1.607 / .144	0	0.685	0
Two On	16	2.358*	1.420 / .144	0	2.039 / 0.779	0.489*
Two On, One Off	8	1.772	1.460* / 0	0.396	1.504 / Var.	0.267*
One On, Two Off	8	1.513	1.624** / 0	0.563	0.550/Var.	0
Three On	8	1.761	1.064 / 0	0.021	1.600/Var.	0
Three On, One Off	4	1.191	Var. Low	0	1.066/ Var.	0.067
Two On, Two Off	4	1.179	Var. Low	0	0.972/Var.	0.156*
One On, Three Off	4	.972	Var. Low	0	.484/Var.	0

Table 7. Table shows the expected reductions in uncertainty (EU) for the top ten interventions that were eventually computable for the six node problems from a flat prior (column 3). Also the EU's and selection frequencies for the same problems were computed using the restricted hypothesis spaces for blocks 3 and 4 respectively as shown in columns 4-7. Interventions had different values depending on *which* colours were fixed. In block 3 any interventions on yellow nodes diminished the EU of an intervention, so the high figures in column 4 are those with only blue nodes fixed and the low have all yellow nodes fixed. Similarly.in block 4, column 6, the high values were for interventions fixes spread over both colours.

Intervention choices in the *last problem faced* in blocks 3 and 4 were investigated for those participants who had been able to describe the roles of the colour cues for those blocks, on the assumption that, adaptations in learning strategy due to the grammar would be apparent by this point. For block 3, even participants who could report the role of the colour cue still fixed yellow 'effect' nodes on 50.0% of their interventions, and the efficient interventions: 'two-on, one-off' and 'one-on, two-off' on the blue nodes only, were only selected 19.12% (and all by one participant). For the last seen problem in block four, 79.8% of interventions were performed within one colour only, leaving the other colour free. This accords with the intuition that participants would tend to focus their attention on one colour at a time once they understood the within-colour-causation-only cue representing a paradigm of divergence of serial human

learning from the parallel learning of the Bayesian active learning model. Based on this limited data it is not really possible to confirm or disconfirm the role of the cue in participants' inferences but only to guide future work.

Replication of results

Efficiency and Sensitivity. As with Experiment 1, participants' interventional efficiency on blocks 1 and 2 were tested and results were very similar. High performers selected the most efficient intervention 25.0% (SD=16.1%) of their tests in blocks 1 and 2, averaging 48.4% (SD=14.5%) of the possible entropy-per-test. Low performers selected the most informative intervention a little more often than in Experiment 1, 17.9% of the time (SD=10.9%) but still averaged 38.7% of the possible entropy per test. This time both high and low performers had a significantly above chance probability of selecting the most efficient intervention ($p<.001$ and $p.=.002$ respectively).

High performing participants endorsed structures when what they had seen justified having a subjective degree of belief of .63 (SD= .19), which is significantly higher than in Experiment 1 ($p<.001$)

Additionally, participants in the high scoring cluster were significantly more likely ($p<.001$) to select the subjectively most likely hypothesis when they had finished performing tests. They did so 60.0% (SD=18.3%) of the time while the mediocre performers did so only 14.0% (SD=13.5%). This measure was not significant in Experiment 1.

Strategies. In Experiment 2, 14 of the 20 participants, including nine of the 10 high performers reported and described a strategy. All 14 reports were, once again, possible to classify either as describing simple endorsement (n=4) or simple endorsement with disambiguation (n=10). There was a marginal positive effect on performance, going from no strategy to strategy ($p=.054$), but no significant difference between scores for simple endorsers

and disambiguators. Once again, purported *simple endorsers* endorsed more wrong links than the *disambiguators*, and this time the difference was significant, 53.25 (SD= 33.925) vs. 19.20 (SD=13.863), p=.026. When this is compared to how many of the 66 *correct* links were also identified (39.5 and 54.1 respectively) this amounts to a very high false positive rate for those who did not perform the disambiguation step. However, simple endorsers in Experiment 2 were *not* more efficient with their interventions choices (43.4% vs. 47.8%), or more likely to select optimal interventions (23.5% vs. 25.3%) than disambiguators as had been the case in Experiment 1.

Collapsed over blocks 1 and 2 of both experiments, there was a clear main effect of strategy: $F(2,36)=12.4$, $p<.001$, with a significant difference going from *no-strategy* to strategy (mean scores block 1 and 2: -74.2 and 131.52, $p<.001$) *and* from *simple endorsement* to *disambiguation* (means: 45.0 and 189.2, p=.018).

The same five models were fit to the participants in Experiment 2 as in Experiment 1[3]. Once again the Bayesian model did not provide the best fit to any participants in blocks 1 and 2. Nine of the 10 high performers were best described (over all four blocks and according to the BIC measure), by the *disambiguation* model with the remaining participant better fit by the *random (repetition)* model. Three low performers were best fit by the *simple endorsement* model, three by the *disambiguation* model and four by the *random (repetitions)* model. Once again, this is broadly confirmatory of the findings in Experiment 1, with a leaning toward the disambiguation model, particularly for successful participants.

Collapsed over both experiments 13 of the 15 purported disambiguators were also best fit by the *disambiguation* model, while only two of the nine purported simple endorsers were best

[3] The links endorsed by participants at interim time points were, unfortunately, not stored properly for Experiment 2, meaning that the model fits were done just using participants' actions.

fit by *simple endorsement*, with the other seven also being better described by the *disambiguation* model. This means that most of those who described only simple endorsement had also been performing at least some disambiguating steps in addition to simply fixing single nodes the whole way through the experiment.

General Discussion

Overall, these experiments show that people can be adept at interacting with systems involving up to six variables, and can often identify or approximate the correct causal structure. Participants' subjective reports allowed us to identify simple active and inferential steps (sequential intervention and disambiguation), and models derived from these steps provided a good fit to participants' actions as well as being effective normative strategies for active causal learning.

The experiments also demonstrated the feasibility of an 'optimal' active learning model, showing that it is a powerful tool for investigating exactly *what* the outcomes of different actions tell you about the structure of the system or environment of interest. This makes it possible to assess the performance of active learning strategies in any well-defined environment. We believe that the ability to quantify active learning actions in this way has potential to improve the understanding of the dynamics of active learning, and also has potential applications in experimental design and artificial intelligence.

Ecological Validity

The patterns of intervention choices by the "optimal" active learning model for the 3-variable problems (Figure 6); and the intervention choices of the disambiguation model (Figure 7), were fundamentally similar, although ordered somewhat differently. However, the intervention choices of the two models diverged increasingly as the number of variables

increased. As the structures in the experiment got more complex, the simple strategies stayed simple, yielding similar performance with only linear increases in required interventions. Meanwhile, the Bayesian solution diverged, selecting more complex interventions as the hypothesis space grew, and fast becoming intractable. These interventions had increasingly obscure effects on the probability distributions over possible structures requiring inordinately more computation. This led to generally better fits for the heuristic strategy models to participants' actions and endorsements than the Bayesian "optimal" learner model.

These experiments do not, however establish whether these strategies are generalizable to application in environments or domains with different causal properties. Noisier environments, with stronger or weaker causal links, and sparser or denser structures all might necessitate changes to the strategies identified here. Similarly, continuous variables, or variations in combination functions are important elements of many causal beliefs, and these are not addressed by the current approaches. However, these strategies essentially break structural induction down into a process of asking simple, targeted, questions of the data. Essentially, they ask of objects: "What does this thing do?" and when faced with several answers, ask: "Are these direct or indirect effects?" Intuitively, this means that that they may work across a broad range of situations. By intervening positively, activating target 'cause' nodes, the precise parameters of a network can largely be ignored, or lumped together as unspecified noise. This avoids the need for exact knowledge about (or at least explicit estimation of) of a causal domain's underlying parameters. While this is something required for the Bayesian model, it is, prima facie, a rare commodity in real-world causal learning.

Essentially strategies are achievable within the processing constraints associated with normal working memory and explicit reasoning (Miller, 1956; Cowan, 2001), while the Bayesian

model quickly becomes intractable. In order to make modest savings in the number of interventions required, the Bayesian model requires inordinately more processing power making Bayesian-intervention-selection a poor choice for balancing the costs and benefits of active causal learning in real environments regardless of what one assumes the computational architecture of the brain or the exact properties of real environments to be. Furthermore, we are typically not completely ignorant about the properties of real environments. Thus, it is uncontroversial that: (i) The actual environment has a highly complex structure, (ii) Processing frugality is at an evolutionary premium in organisms with limited energetic resources, and (iii) Active exploration of (at least the local and physical) structure of our environment is a central part of being an embodied cognitive agent living in the world (Simon, 1956; Clark, 1998). When structure is complex, mechanisms that scale up linearly are better than those that suffer from combinatorial processing explosions. And, when active intervention is cheap, maximising your interventional efficiency becomes secondary to having robust and efficient mechanisms for active causal structure induction.

To follow up from these experiments, it would be instructive to explore how different learning strategies perform in different environments, and how easily people can adapt or find an efficient active learning strategy moving between environments with different causal properties. Investigating this could guide hypotheses about whether we are best described as having a single adaptable, or multiple selectable active learning mechanisms or if we merely generate good enough strategies 'on the fly'. Perhaps we are able to use our implicit perceptual sensitivity to covarational frequencies and patterns (Lewicki, Hill & Czyzewska, 1992) to continually readjust the rigor of our action selection and causal ascription mechanisms as we attend to different aspects of our surroundings.

Causal Attribution and Individual Differences

One caveat to these results comes in the form of the large individual differences between performances in the task. A significant number of participants over both experiments performed very poorly (18/39). We hypothesise that this is partly due to the weakness of the task interface as some participants reported struggling with adding and removing arrows when several nodes were close together, or finding the visualisation of the fixed nodes unintuitive. While it is likely there are individual differences in causal learning, causal information is so fundamental to cognition that it would be implausible to claim many people lack basic active causal learning ability. However, the very good performance of the other participants in what was a highly demanding task shows that we do have these causal learning abilities and are able to identify much more complex causal structures than participants are usually given credit for in the causal learning literature.

Additionally, comparison between the performance of the simple link endorsement and the disambiguation strategies suggests an interesting potential divergence between basic causal attribution mechanisms. Prima facie, much of the role of causal knowledge is pragmatic. We often use our causal knowledge to get answers to questions like: "What do I need to do in order to make X occur?" and are often happy with approximate and simplified causal representations for situations which we know to be more complex when more closely investigated (Casini et al, 2011). In the light of these considerations, the simple link endorsement strategy, which does not worry about the possibility of causes being indirect, seems to be a pragmatically sensible approach to endorsing causal structure. However, it is often vital to establish, not just the pragmatic, but also the unique, counterfactually reliable structure of a domain. For example, to establish whether a new drug really works, or if an expensive policy will lower crime, one must

control for the possible indirect action of placebo or policy change effects, meaning that in many situations additional disambiguation steps can be very important.

A possibility is that there are individual differences, specifically in peoples' propensity to disambiguate between multiple potential causal paths from actions to events which might go some way to explaining deep ideological differences between individuals. For example, people differ markedly in their propensity to endorse "overdetermined" (Lewis, 1973) explanations of worldly phenomena. Examples include believing in direct magical, supernatural or religious connections between actions and events, in addition to their established indirect physical-causal explanations. It might be that some individuals are more inclined to accept correlations between actions and its distal effects as direct cause-effect relations, in the same way that *simple endorsement*, endorses direct links to everything that follows from an action. Thus, individual differences in basic active learning mechanisms might determine how individuals develop their systems of causal beliefs.

Cue Integration

Finally, the colour cue in Experiment 2 provided some evidence that people can learn grammatical level cues over multiple learning instances, and that they benefit from this knowledge in subsequent learning episodes. However, it was not possible to investigate this effect in detail because of the relatively small number of problems within a block, and the relatively small number of successful participants who learned the cue. However, participants' responses provided some useful information to guide follow up studies. For example, some participants reported paying no attention to the colours at all. It may be that the task put such a load on attention and working memory that many participants barely noticed the colour of the nodes they were testing. Two participants reported using them as placeholders for remembering which nodes they had already tested, suggesting the interpretation that learning was focused only

at a single 'level' at a time. Potentially, noticing the role of the colours might have been dependent on whether participants focused their attention directly on them, temporarily ignoring the primary goal of the task, or learned the relations when 'resting' on the feedback screens between problems. A follow up experiment, with more tests and no feedback phase, is necessary to really get to the bottom of this. Another observation was drawn from participants' verbal reports and the fact that more people learned the colour cue in only one block (7) than both (3). This was that realising that blues cause yellows in block 3 may have had a detrimental effect on participants' chances of realising that greens cause greens and oranges cause oranges in block 4. This coheres with the hierarchical framework considering that block 3 constituted all of the experience that participants would have had with the causal properties of 'coloured nodes'. Participants would be perfectly sensible in having an expectation at the start of block 4 that colours have something to do with the causal roles of nodes (e.g. as either cause or effect) making it unlikely that they would consider new colours to correspond to the, conceptually very different, causal "groupings" of block 4.

Conclusions

Overall, these results show that people can learn complex probabilistic causal structure, without domain knowledge, so long as they are given a free rein to intervene. In the experiments reported here, people achieved this by sequentially targeting interventions on single objects, finding out what they do, and building up a model on the basis of this, often performing additional 'disambiguation' steps when what they saw could correspond to multiple linkages. The intervention choices of the simple strategies identified in this paper correspond closely to the Bayes-'optimal' intervention choices for simple problems, proving that this is an efficient way to work on a local scale, but the strategies also scale up in a way that Bayesian active learning does not.

The results are in line with the rational hierarchical framework, in that people show some signs of being able to identify and make use of a minimal 'grammars' from simple binary cues, although follow-up studies are required to show this more forcefully. Additionally, the effort required in completion of the task, and the difficulty many people found in solving it, testifies to the extent to which people normally rely on their existing knowledge in making causal inferences, only learning actively to the degree that a situation is unfamiliar.

References

Akaike, H. (1974). A new look at the statistical model identification. In IEEE Transactions on Automatic Control 19 (6), 716–723.

Bromberger, S. (1966). The Journal of Philosophy, 63 (20), 597-606.

Casini, L, Illari, P. M., Russo, F & Williamson, J. (2011). Models for prediction, explanation and control: recursive Bayesian networks. Theoria, 26, 1, 5-33.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. Psychological Review, 104, 367-405.

Chomsky, N (1975). Reflections on Language. New York: Pantheon.

Clark, A. (1998). Being There. Cambridge MA: MIT Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences, 24, 87–185.

Sloman, S. A., Fernbach, P., Ewing, S. (2009). Causal models: The representational infrastructure for moral judgment. In B. H. Ross (Series Ed.) & D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Eds.), Psychology of Learning and Motivation, 50: Moral Judgment and Decision Making. San Diego, CA: Academic Press.

Gilovich, T., Griffen, D. & Kahneman, D. (2002). Heuristics and Biases. Cambridge: Cambridge University Press.

Glymour, C. (2001). The mind's arrows: Bayes nets and graphical causal models in psychology. Cambridge, MA: 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, Glymour, C, Sobel, D, Schulz, L, Kushnir, T & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. Psychological Review, 111, 1-31.

Griffiths, T. L. & Tenenbaum, J. B. (2009). Theory-based causal induction. Psychological Review, 116, 661–716.

Hagmayer, Y. & Lagnado, D. A. (Forthcoming). Causal models in judgment and decision making. In Dhami, M., Schlottman, A., Waldmann, M. (ed.) Judgment and decision making as a skill: Learning, Development and Evolution. Cambridge: Cambridge University Press.

Hagmayer, Y., Sloman, S., Lagnado, D., & Waldmann, M. R. (2007). Causal Reasoning Through Intervention, 86-101.

Hansen, M. & Yu, B. (2001). Model selection and the principle of minimum description length. Journal of the American Statistical Association, 96, 454, 746-774.

Holyoak, K. J. & Cheng, P. W. (2011). Causal Learning and Inference as a Rational Process: The New Synthesis. Annual Review of Psychology, 62, 135–63.

Jones, M & Love, B. C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. Behavioural and Brain Sciences (2011).

Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgment Under Uncertainty: Heuristics and Biases. New York: Cambridge University Press.

Kant, I. (1781/1965). Critique of Pure Reason. London: Macmillan.

Kemp, C., Goodman, N. D., & Tenenbaum, J. B. (2010). Learning to Learn Causal Models. Cognitive Science, (2011).

Krynski, T. R. & Tenenbaum, J. (2003). The role of causal models in reasoning under uncertainty. In Proceedings of the 25th Annual Conference of the Cognitive Science Society. Mahwah, NJ: Erlbaum.

Kushnir, T, Gopnik, A, Schulz, L. E. & Danks, D. (2003). Inferring hidden causes. In R. Alterman & D. Kirsh (Eds.) Proceedings of the Twenty-Fifth Annual Meeting of the Cognitive Science Society, (pp. 699–703). Mahwah, NJ: Erlbaum.

Lagnado, D. (2011). Causal thinking. In McKay-Illari, P., Russo, F., Williamson, J. (ed.) Causality in the Sciences. Oxford: OUP.

Lagnado, D. & Sloman, S. (2002). Learning causal structure. In W. Gray & C. D. Schunn (Eds.), Proceedings of the twenty-fourth annual conference of the cognitive science society, Mahwah, NJ: Erlbaum.

Lagnado, D. & Sloman, S. (2004). The advantage of timely intervention. Journal of Experimental Psychology: Learning, Memory & Cognition, 30, 856-876.

Lagnado, D., & Sloman, S. (2006). Time as a Guide to Cause. Cognition, 32 (3), 451- 460.

Lagnado, D., & Speekenbrink, M. (2010). The influence of delays in real-time causal learning. The Open Psychology Journal, 3(2), 184-195.

Lagnado, D., Waldmann M., Hagmayer, Y & Sloman, S. (2007). Beyond covariation: cues to causal structure. In A. Gopnik, A. & L. Schulz (Eds.), Causal Learning: Psychology, Philosophy, and Computation, 154–72. London: Oxford University Press.

Laplace, P. L. (1814/2009) Essai philosophique sur les probabilities. New York: Cambridge University Press.

Lewicki, P., Hill, T., Czyzewska, M. (1992). Nonconscious acquisition of information. American Psychologist, 47, 6, 796-801.

Lewis, D. (1973). Causation. The Journal of Philosophy, 70, 17, 556-567.

Marr, D (1982). Vision. New York: Freeman & Co.

McCarthy, J. & Hayes, P.J. (1969), Some Philosophical Problems from the Standpoint of Artificial Intelligence. In D. Michie & B. Meltzer (Eds.) Machine Intelligence. Edinburgh University Press.

Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 63, 2, 81-97.

Myung, I. J., Forster, M. & Browne, M. W. (2000). Tutorial on maximum likelihood estimation. Journal of Mathematical Psychology, 44, 1–231.

Nyberg, E. & Korb, K. (2006). Informative Interventions. In F. Russo & J. Williamson (Eds.) Causality and Probability in the Sciences. College Publications: London.

Pearl, J. (2000). Causality. New York: Cambridge University Press (2nd edition).

Scheines, R., Easterday, M. & Danks, D. (2007). Teaching the Normative Theory of Causal Reasoning. *Department of Philosophy,* 426, http://repository.cmu.edu/philosophy/42.

Simon, H. (1956). Rational Choice and the Structure of the Environment. Psychological review, 63, 129-38.

Shanks, D. R. (2007). Forward and backward blocking in human contingency judgement. The Quarterly Journal of Experimental Psychology Section, 37, 1, 1-21.

Shannon, C. E. (1951). Prediction and entropy of printed English. The Bell System Technical Journal, 30, 50-64.

Sloman, S. (2005). Causal models: How people think about the world and its alternatives. Oxford: OUP.

Sloman, S. & Lagnado, D. (2005). Do we "do"? Cognitive Science, 29 (1), 5-39.

Sprites, P., Glymour, C. & Scheines, R. (1993/2000). Causation, Prediction, and Search (Springer Lecture Notes in Statistics). Cambridge, MA: MIT Press. (2nd edition).

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E., & Blum, B. (2003). Inferring causal networks from observations and interventions. Cognitive Science, 27, 453-489.

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. Trends in Cognitive Science, 10, 309-318.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., and Goodman, N. D. (2011). How to Grow a Mind: Statistics, Structure, and Abstraction. Science 331, (6022), 1279-1285.

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.

Woodward, J. (2003). Making things happen: A theory of causal explanation. Oxford: Oxford University Press.