Active causal structure learning in continuous time

Tianwei Gong
University of Edinburgh

Tobias Gerstenberg
Stanford University

Ralf Mayrhofer
University of Göttingen

Neil R. Bramley
University of Edinburgh

Author Note

A preliminary version of the ideal observer model was presented at the 39th Annual Meeting of the Cognitive Science Society and appeared in the non-archival conference proceedings (Bramley, Mayrhofer et al., 2017). TiG is supported by the Edinburgh University PPLS Scholarship. RM is supported by a Deutsche Forschungsgemeinschaft Research Grant (WA 621/24-1). NB is part supported by a EPSRC New Investigator Grant (EP/T033967/1).

Abstract

Research on causal learning has largely focused on learning and reasoning about contingency data aggregated across discrete observations or experiments. However, this setting represents only the tip of the causal cognition iceberg. A more general problem lurking beneath is that of learning the latent causal structure that connects events and actions as they unfold in continuous time. In this paper, we examine how people actively learn about causal structure in a continuous time setting, focusing on when and where people intervene on the system and how this shapes their learning. Across two experiments, we find that participants' accuracy depends on both the informativeness and evidential complexity of the data they generate. Moreover, intervention choices strike a balance between maximizing expected information and minimizing expected inferential complexity. That is, we find people time and target their interventions to create simple yet informative causal dynamics. We discuss how the continuous-time setting challenges existing computational accounts of active causal learning, and argue that metacognitive awareness of one's inferential limitations plays a critical role for successful learning in the wild.

Keywords: causal learning, intervention, time, resource rationality, causal cycles

Active causal structure learning in continuous time

Introduction

The ability to predict, plan, and control events in the world demands a sophisticated representation of the world's causal structure. Learning such a causal model requires gathering causal evidence that comes from *interventions* (Pearl, 2000) — actions that manipulate the environment in ways that reveal causal direction and distinguish spurious correlations from genuine relationships. However, learning causal structure in general, and selecting interventions in particular, are computationally challenging problems even under idealized conditions (Bramley, Dayan et al., 2017). In everyday life, this challenge is compounded by the need to interact with the causal environment in real time, bringing computational and bandwidth constraints to the fore (Griffiths et al., 2015; Simon, 1982). In this paper, we explore how people actively learn about causal structure in real time. To do this, we introduce a causal learning task in which participants interact with causal devices in real time, deciding when and where to intervene in order to gather information. To motivate our novel experiments and modeling, we first summarize prior empirical work on active causal learning and point out some of its limitations. We then introduce notions of resource-rational behavior (Lieder & Griffiths, 2020) that serve as a guideline for our computational modeling before turning to our experiments, model fits and their implications.

Prior work on active causal learning

Everyday cognition is rich with causal beliefs that explain the progression of events, shape our predictions about what is to come, and allow us to choose actions to realize our goals. For example, you might recognize a squeaking sound as caused by the opening of your garden gate, predict the doorbell will ring with your takeaway and get up to answer the door in anticipation. Many researchers have used a causal Bayesian network framework to study how people build up and represent networks of causal mechanisms and affordances (Bramley, Dayan et al., 2017; Griffiths & Tenenbaum, 2009; Lagnado & Sloman, 2002; Lucas & Griffiths, 2010; Meder et al., 2010; Rehder, 2014; Rottman & Hastie, 2016; Schulz et al., 2007; Sobel & Kushnir, 2006; Stephan et al., 2021; Steyvers et al., 2003). While the particulars of these studies are diverse, many share a core set of properties illustrated in Figure 1a. Participants are typically asked to distinguish between a set of candidate causal structures on the basis of evidence. Often this evidence takes the form of "snapshot" samples of discrete variables' states. Most often, the variables of interest have been taken to be binary with one value construed as a variable being "present" or "active", and the other state as "absent", or "inactive".

In the covariation-data setting, observational samples are insufficient to uniquely reveal structure (Pearl, 2000; Spirtes et al., 2000). For example, if a learner observes two variables co-occurring, they cannot tell if one is causing the other or if they share some unobserved common cause. One solution is to *intervene* (Pearl, 2000) — manipulating one or more variables in the system by fixing them to particular values and observing how this effects the rest of the system. A number of experiments have allowed participants to choose which variable or variables to fix through intervention, essentially choosing what sequence of experiments to perform in order to support their learning (Bramley, Dayan et al., 2017; Bramley et al., 2015; Coenen et al., 2015; Lagnado & Sloman, 2002; Steyvers et al., 2003).

Well chosen interventions can speed up learning, allowing learners to target their uncertainty and narrow in quickly on the true model. However, poorly chosen interventions can be worse than random actions or passive observations by perseverating on the wrong parts of the data space and failing to distinguish between key possibilities (Settles, 2009). In the covariation-data setting, research has shown that adults and children are able to select informative interventions and learn successfully from them about probabilistic systems involving a handful of variables (Bramley et al., 2015; Coenen et al., 2015; McCormack et al., 2016; Meng et al., 2018; Steyvers et al., 2003). At a normative level, informative interventions are those whose consequences are expected to strongly distinguish among the potential hypotheses, maximally decreasing global uncertainty in expectation (Tong & Koller, 2001), or maximizing the chances of holding the correct belief or one that serves one's goals (Nelson, 2005). A number of experiments have demonstrated broad alignment with these norms in both adults and children, but also departures that hint at process-level considerations (Bramley, Dayan et al., 2017; Bramley et al., 2015). For example, choices that are expected to confirm or refute a currently-favored hypothesis are often selected over those that provide more information about the full hypothesis space (Coenen et al., 2015; Klayman & Ha, 1989; Meng et al., 2018; Steyvers et al., 2003). Intervention strategies also sometimes reflect application of generic strategies such as systematically fixing the values of some variables while varying others in order to isolate one potential relationship at a time (Bramley et al., 2015; McCormack et al., 2016; Schulz et al., 2007). One manifestation of this is the so-called *control of variables* strategy in which a set of candidate causal variables are fixed and one variable is changed in each experiment (Chen & Klahr, 1999; Kuhn & Brannock, 1977; Zimmerman, 2007). Following such a strategy has been emphasized in developmental psychology as a marker of mature scientific experimentation, but also shown to be suboptimal in certain environments (Coenen, Ruggeri et al., 2019). Finally, adults' intervention strategies have been shown to be adaptive, depending critically on environmental factors including time pressure and whether a strategy was informative in the past (Coenen et al., 2015).

What prior work has neglected

While previous investigations of active causal learning have been fruitful, they have largely focused on situations that mimic a particularly idealized kind of laboratory conditions. By this, we mean that participants make a sequence of observations or experiments in which interventions are selected and the consequent values of the other

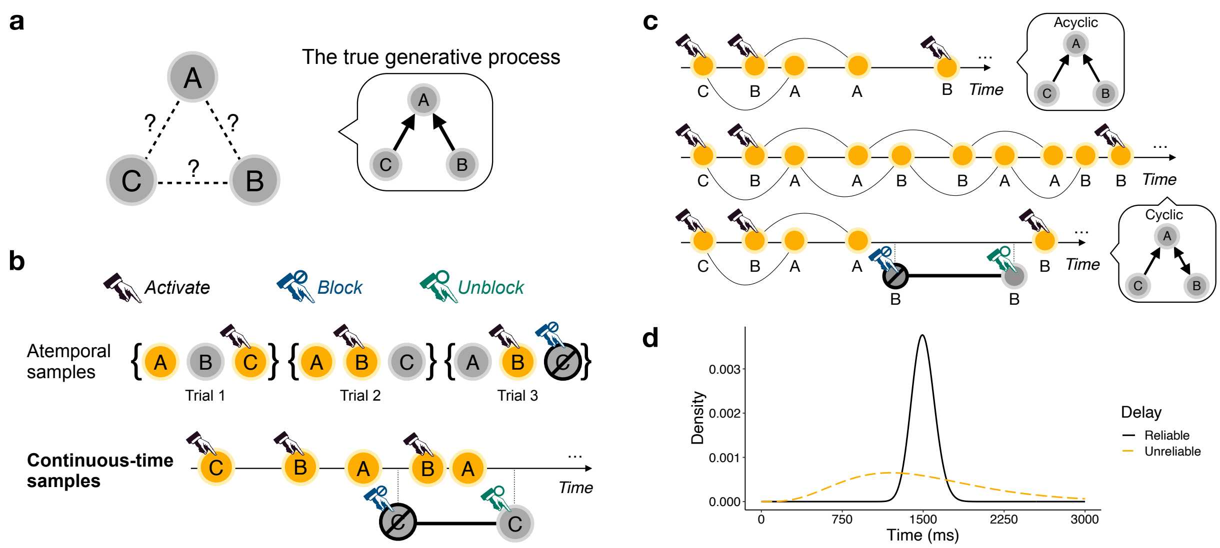

Figure 1

Illustration of causal systems and sample types. a) Three components with causal connections unknown to the learner. b) Atemporal samples under three trials, and its corresponding continuous time samples. Yellow indicates a component activated. In the continuous-time setting, interventions activate components in real time and effects occur intermingled on the timeline. c) Three examples of event dynamics in acyclic structures, cyclic structures, and cyclic structures with blocking interventions. Small curves indicated the underlying generative process unbeknownst to the learner. d) Gamma density distributions under reliable vs. unreliable delays.

variables are revealed all at once. In this way, participants are invited to generate and reason from a series of independent observations and the system of interest is taken to be stateless in the sense that there is no carry-over of variable state from one experiment or observation to the next. Figure 1b illustrates evidence that was generated from interventions and subsequent observations of the variables in the system. At a computational level, the problem is one of identifying the true generative causal Bayesian network — the parameterized graph that captures the patterns of covariation between the variables under both observations and any hypothetical intervention (Pearl, 2000). For example, in Trial 1 in Figure 1b we see that, conditional on an intervention that activates C, A activated and B did not. This can be written as $\{A = 1, B = 0 | \,\mathrm{Do}[C = 1]\}$, where 1 indicates a variable was active, 0 indicates it was inactive and Do[.] indicates a variable was fixed through intervention and thereby temporarily disconnected from its normal causes. Interventions can involve multiple variables. For example, in Trial 3, both B and C are manipulated $\{A = 0 | \,\mathrm{Do}[B = 1, C = 0]\}$. In this kind of task, ideal inference and intervention selection are well understood computationally, facilitating comparison between behavior and rational norms (Rottman & Hastie, 2014). However, we argue that this task setup differs in several respects from the causal learning and reasoning problems people face in daily life: (1) the consideration of time, (2) interdependence, and (3) cycles.

Time

Temporal information, including the order of events and the delays between them, was generally not provided in previous studies. For example, Coenen et al. (2015) described a cover story of computer-chip systems with causal relationships as the passage of electrical current between the components, occurring too fast to distinguish order of activation. Other studies only allowed participants to view the final outcome (Bramley et al., 2015; Rottman & Keil, 2012). In contrast, many everyday causal relationships take time to propagate, meaning that the temporal order and delay between events is relevant for both inferring causal relationships and to the kinds of predictions that the resulting models support. Correspondingly, the notion that causes must precede their effects is foundational to the concept of causation (Burns & McCormack, 2009; Lagnado et al., 2007; White, 2006). Indeed, people have been shown to rely on temporal order to guide causal inference even when it conflicts with covariation information (Lagnado & Sloman, 2006) and to assign low probabilities to mechanistic explanations for event sequences that imply an effect occurring at the same time as its cause (Bramley, Gerstenberg, Mayrhofer et al., 2018). People not only have expectations about the order of events but also about the delays between them, giving higher causal strength ratings when delays between a putative cause and effect are short and reliable (Greville & Buehner, 2010) as well when they conform to prior or mechanistic expectations (Hagmayer & Waldmann, 2002). For example, Buehner and May (2002) found that participants gave higher causal judgments about a switch turning on an energy-saving light bulb if the bulb came on a few seconds after the switch was pressed rather than instantly (consistent with the behavior of early energy saving bulbs). As such, temporal information has received separate attention in causal learning studies and is yet to be examined in the context of intervention selection.

Interdependence

Under the laboratory conditions created in prior experiments, evidence is taken to come from multiple “independent, identically distributed” (i.i.d.) observations or interventions. For example, trials may pertain to different test subjects drawn from the same population (e.g. pairs of patients and treatments; Buehner et al., 2003), or might involve repeated interactions with the same causal mechanism, but collected via a protocol that ensures variables “reset” from one trial to the next (e.g. the “blicket detector”; Gopnik et al., 2001; Lucas et al., 2014). However, it is rare for everyday experience to exhibit these properties. In life, there is no magic reset button. It is hard to be sure whether and when a causal system has been reset without understanding its underlying mechanism (defeating the goal of the exercise). To illustrate this, imagine noticing your puppy is unusually excited one evening and wondering why. Perhaps his elevated mood is due to a new variety of dog food you fed him at 5pm, perhaps it is because of a new floral scent on the road where you walked him at 6pm. The puppy might still be happy about his dinner even after smelling the flowers. A poor approach to resolving the question would be to always feed him beside the flower bed. Better might be to vary the relative time of walking and feeding him while keeping a close eye on the time intervals implied by different causal explanations. This example illustrates that active learning in our everyday lives is better understood as a rolling sequence of interventions, with cause and effect events unfolding on a single continuous timeline. In fact, Rottman and Keil (2012) found that when presented with a sequence of experimental results, even paired with a cover story that implied these experiments were independent, participants frequently did not judge causal relationships according to the contingencies that arise from treating the samples as i.i.d, but rather by how values changed relative to their state on the previous experiment (see also Derringer & Rottman, 2018). This suggests that when evidence arrives over time, people default strongly to a presumption of temporal dependence. Thus, it seems that temporal dependence not only reflects genuine causal phenomena but that it may also better match laypeople's intuitive causal theories than time-agnostic Bayesian networks do.

Causal cycles

Causal learning studies have largely focused on acyclic causal systems where causal influences flow only in one direction, never revisiting the same components. This is partly due to the conceptual and mathematical convenience afforded by acyclic causal Bayesian network formalism (see Rottman & Hastie, 2014, for a review). However, the continuous-time setting enables us to investigate cyclic causal relationships, as effects of reciprocal components can occur in alternating fashion (Figure 1c). This is important because many causal processes in the natural world are cyclic (Malthus, 1872), and people also frequently report causal beliefs that include cyclic relationships when allowed to do so in experiments (Kim & Ahn, 2002; Nikolic & Lagnado, 2015; Rehder, 2017; Sloman et al., 1998).

The current paradigm

We use a different paradigm than past work to study causal learning outside of laboratory settings. In our task, there are no independent trials. Instead, actions and events occur during a single continuous stream of time. We thus represent causes and effects as point events and assume that effects follow their causes with some, stochastic but predictable, delay. In this setting, the timing of causally unrelated events are temporally "untethered" to one another such that the timing of one provides no information about when to expect the other, while the presence of direct or indirect causal influence implies there will be temporal dependence and predictability. The point-event representation has been used in a number of recent studies of temporal causal reasoning. Like the contingency data setting, it rests on a firm mathematical foundation that supports normative inferences from temporal information to causal structure. Greville and Buehner (2007), Griffiths and Tenenbaum (2009) and Pacer and Griffiths (2012, 2015) first demonstrated that people can infer how pairs of variables affect one another from sequences of point events. Pacer and

Griffiths' (2012, 2015) modeling showed these inferences reflect dependence between the occurrence of putative cause events and the *rate* at which the relevant effect events occur over time. Some recent research has focused on so-called "actual causal inference" (Halpern, 2016) — from a known causal structure to judgments about which events actually caused one another in a given observation (Stephan et al., 2020). Recent work has also focused on the general problem of structure induction from temporal information via reasoning explicitly about the potential actual causation pathways (Bramley, Gerstenberg, Mayrhofer et al., 2018; Gong & Bramley, 2020; Valentin et al., 2020). For example, Bramley, Gerstenberg, Mayrhofer et al. (2018) explored how people make structure judgments based on one or several episodes in which components of a causal system activated over time. Participants' judgments were best predicted by inference to the best model under a combination of hard order constraints and soft delay expectations. Even when order information was fixed — e.g. components A, B and C all activated in the same order on every episode — participants were still sensitive to the variation in inter-activation delays and distinguished if the data had been generated by an $A \rightarrow B \rightarrow C$ chain or a $B \leftarrow A \rightarrow C$ fork.

While time was central to the evidence in the above studies, the data was not fully continuous in the second sense of lacking independent trials. Participants observed several videos of the same causal system on different occasions. In order to make things happen, each clip began with the system at rest perturbed by an exogenously caused root-component activation, with effects following from there. Since components could only activate once in these tasks, the system would quickly reach a steady state. This still departs from a fully continuous-time setting in which interventions and effects are intermingled and components may exhibit multiple activations within the same episode. In the fully continuous setting, there is more complexity and ambiguity about what caused what since each effect event might be attributed to any earlier-occurring event and might have its own effects that are still to occur.

This paper aims to investigate adults' performance in a causal learning task that incorporates both temporal information and pervasive interdependence, and includes both acyclic and cyclic causal systems. We develop the first computational account of real time inference and efficient intervention choice and use this to explore human behavior in our experiments.

Cognitive resource limitations

Minds are bounded in their capacity to compute and store information which means that human reasoning and decision-making deviates from computational-level norms and ideals (Anderson, 1990; Simon, 1982). Recent framings of bounded rationality have argued that human minds have evolved or discovered various heuristics and strategies that trade off efficiently between the costs of computation and the rewards of greater accuracy (see Griffiths et al., 2015; Lieder & Griffiths, 2020; Shenhav et al., 2017, for review). These *resource-rational analyses* have argued that a number of decision making phenomena traditionally seen as irrational — including as anchoring and probability matching — may instead represent efficient solutions to a computation–value trade-off under some sensible approximation scheme (Dasgupta et al., 2017; Hawkins et al., 2021; Lieder et al., 2018). Given the computational complexity and conceptual centrality of structure learning, we expect computational costs weigh heavily in this trade-off when structure inference must take place in real time (Christiansen & Chater, 2016). As an analogy, in the idealized laboratory setting, causal reasoning is a little like crafting an essay: evidence can be collected, organized and put together carefully with room for reorganizing and backtracking in searching for an effective structure. However, real-time learning more closely resembles the problem of writing under exam conditions: one must react immediately to the question bringing ones inferential tools to bear quickly and efficiently without the luxury of time to backtrack.

Humans can only consciously process a limited about of information at any given time (Gilbert, 1966; Lai & Gershman, 2021; Miller, 1956) and this leads to what has been called a "now-or-never bottleneck" on learning (Christiansen & Chater, 2016). As an example, an average human's ability to parse natural language input tops out around 60 bits per second (Gilbert, 1966; Pierce & Karlin, 1957; Verghese & Pelli, 1992). Critically, in self-paced reading tasks, participants will slow down when they encounter surprising inputs that require deeper processing (Frank, 2013; Lowder et al., 2018). Indeed, reaction time is taken to be a foundational measure of explicit processing, predicated on the assumption that our processing speed is broadly constant (Galton, 1890). We predict that similar meta-learning considerations will shape real-time active causal learning since, like language comprehension, it also requires abstraction to symbolic model-based semantics from a continuous input in continuous time. That is, we hypothesize that, in the wild, we intervene on the world, not just to generate information but also to *time* it so that it can be processed effectively.

We apply a bounded rationality framework to our analysis of human active causal induction. This will involve first considering the impact of both information and complexity on inferential success, and second, modeling intervention selection as driven by the goal to maximize expected epistemic utility while minimizing expected computational load. The rest of this paper is structured as follows: First, we introduce our task and lay out our computational-level modeling framework. Second, we report on two experiments that probe continuous time active causal learning under a range of causal structures and delay settings. Third, we lay out our resource-rational account for participants' behaviors our experimental conditions.

Learning environment

We explore active learning about abstract causal "devices" that are made up of three to four components connected by causal relationships (Figure 1a). The causal relationships produce point events, visualized as activations of the device's components over time (e.g. Figure 1c). For causally related components, an activated component will usually activate each of its effect components once after some parametric delay (described below). Each event can then trigger subsequent events according to its causal connections and so on.

We focus on systems in which components never spontaneously activate but where the causal relations are not perfectly reliable. Concretely, we assume a success probability of $w = 0.9$ for all causal connections in all devices in our experiments. This means that when there is a causal link from A to B and A is activated, then there is a 90% chance that B will activate because of A after some delay. For example, Figure 1c shows a possible timeline of events in which a learner intervenes on component C, then on B, then on B again. When the true structure (unknown to the learner) is a $B \rightarrow A \leftarrow C$ (so called "collider" or "common effect"), two activations of A follow the interventions but with some ambiguity about their actual source. In the causal devices we test, any pair of components might be unconnected or otherwise have a directed ($A \rightarrow B$ or $B \rightarrow A$) or bidirectional ($A \leftrightarrow B$) causal connection. This results in a hypothesis space of 64 possible 3-variable and 4096 possible 4-variable devices.

Delays

Following recent studies of observational causal inference from temporal information, we use gamma distributions to model causal delays (Bramley, Gerstenberg, Mayrhofer et al., 2018; Stephan et al., 2020; Valentin et al., 2020). Gamma distributions define a density over $(0, +\infty)$ natively using a shape α and rate β parameter that allows for a variety of delay shapes with differing means and more or less variability. In our experiments, we contrast causal systems with reliable delays ($M \pm SD = 1500 \pm 100$ ms, i.e. $\alpha = 200, \beta = \alpha/1500$, see Figure 1d) with systems with unreliable delays (1500 ± 700 ms, i.e. $\alpha = 5, \beta = \alpha/1500$).[1] For simplicity, and in order to focus on other aspects of inference, we train participants on the true delay parameters in the instructions and assume they are known to our models.

Interventions

We allow learners to intervene on the causal system in two ways:

1. By *activating* components (Experiment 1 and 2), thus potentially setting in motion a new sequence of events.

[1] $\alpha = 1$ is an important limiting case of the gamma distribution, an exponential distribution in which the expectation has no central tendency around its average, here meaning the expected delay remains constant, no matter how long you have already waited. This can be used to model causal independence, where the occurrence of one event provides no information about when to expect another event (Gong & Bramley, 2020).

2. By *blocking* components (Experiment 2 only), preventing that component from either being activated and from activating any other components until it is unblocked again.

Activating and blocking are superficially analogous to fixing variables to be on $Do[X = 1]$ or off $Do[X = 0]$ in the idealized laboratory setting. However, they also differ in important ways. In the continuous time setting, activating does not disconnect a component from its normal causes. The intervened-on component can be activated again an arbitrary number of times during the same episode either by the intervener or when caused by other variables in the system. For instance, if the activated component is part of a cycle we would expect it to be re-activated repeatedly following its initial activation until one of the causal connections fails (see Figure 1c). Thus, activation is better thought of as a shock to the system than as a form of graph surgery.

On the other hand, *blocking* actions do exemplify the "graph surgery" property in Pearl's (2000) sense. They disconnect the blocked component from its normal causes until it is unblocked again (Figure 1c). In the contingency-data setting, blocking is essential for discriminating certain structures. For example, turning on any single component (i.e. $Do[A = 1]$, $Do[B = 1]$ or $Do[C = 1]$) will generate an identical pattern of dependence under a $A \rightarrow B \rightarrow C$ chain and a $C \leftarrow A \rightarrow B \rightarrow C$ fully connected structure, making it impossible to distinguish these structures from single-variable interventions. To identify whether there is direct link between A and C, one must turn on A while simultaneously blocking or disabling B (i.e. $Do[A = 1, B = 0]$). The few studies that have allowed participants to perform multiple interventions have shown that such "controlled tests" are associated with higher learning accuracy in adults (Bramley et al., 2015), while children appear to have more difficulty using these kinds of tests appropriately (McCormack et al., 2016; but see Schulz et al., 2007). A similar principle to this conditional blocking pattern, known as "control of variables" has been emphasized in educational contexts as a marker of good scientific reasoning (Chen & Klahr, 1999; Kuhn & Brannock, 1977; Zimmerman, 2007). In designing scientific experiments, one typically seeks to either fix or randomize nuisance variables and factors. Both are conceptually related to blocking here in that they help to isolate causal influences (Kuhn, 1989).

The current continuous-time setting endows blocking with different implications. Since causes generate effects individually, blocking is not strictly required to distinguish direct and indirect paths. Going back to the chain vs. fully connected example above, a fully connected model would normally produce two staggered activations of C following an intervention on A while the chain would produce only one, making them distinguishable in principle. Nevertheless, blocking may be a useful intervention so as to reduce computational complexity and ambiguity. If a learner activates a component while some of the other components are blocked, fewer effects will be generated and hence fewer interpretations of the evidence will need to be considered. The remaining events could still be informative and less ambiguous for revealing sub-parts of the overall structure. Well-timed blocks might also be used to impose pseudo-independence and trial-like structure within a continuous interaction. For example, one might block, wait, then unblock components to help reset a system and stop any ongoing activity from complicating the inference process going forward. This could be particularly useful for bounded learners dealing with complex dynamics. In this paper, we contrast a setting where learners can only perform activating interventions (Experiment 1) with one in which both activations and blocks can be used (Experiment 2).

Incentive structure

In order to incentivize efficient testing, we allow each learner a limited budget of interventions — concretely, 6 activations per problem — and a limited total trial length — 45 seconds. Furthermore, we incentive participants to mark the correct causal links as quickly as possible within the trial by rewarding them based on their accuracy at an unknown random time point during each problem. This means it is in their interest to register their best guess accurately and early and update it whenever their conclusions change during a learning episode (cf. Bramley, Dayan et al., 2017). In the modeling the task, the value of future information is thus temporally discounted in line with the linearly increasing probability that the bonus has already been paid. This incentive structure forces an ideal intervener to balance the benefits of waiting until they know more — e.g. to make best use of their next intervention — against the opportunity cost of waiting too long (e.g. and missing their opportunity to use the outcome of the next intervention to improve their model and bonus). Thus, in order to gather evidence efficiently, learners need to not only consider *how* and *where* to intervene next, but also *when* to do so. Further complicating the consideration of when to act for a bounded agent is the issue of managing the expected complexity of the resultant dynamics, and consequential computational cost of unraveling the evidence they provide.

Modeling framework

Building a computational account of bounded active intervention selection involves several interrelated facets:

1. An account of inference from continuous-time causal evidence resulting from interventions.

2. An account of the cognitive costs of performing inference from different patterns of evidence.

3. An account of intervention selection based on expected value of interventions for the learner in terms of both their expected epistemic benefits and expected computational costs.

Ideal observational (IO) learning

We first focus on describing how a learner should ideally update their beliefs after seeing evidence produced from interventions. The ideal observer infers a posterior distribution $P(S|\mathbf{d};\mathbf{i})$ over causal structures $s \in S$ based on evidence $\mathbf{d}$ conditional on interventions $\mathbf{i}$ using the Bayes rule:

$$P(S|\mathbf{d};\mathbf{i}) \propto p(\mathbf{d}|S;\mathbf{i}) \cdot P(S). \tag{1}$$

Here, $P(S)$ denotes the prior probability distribution over causal structures, and $p(\mathbf{d}|S;\mathbf{i})$ denotes the likelihood of the observed data conditional on the interventions under each possible causal structure.

We assume that data $\mathbf{d}$ consists of all non-interventional activation events and their timings indexed by their chronological order and subscripted by the component at which they occur $d^{(\text{index})}_{\text{component}}$ and that this is conditioned on the set of interventions $\mathbf{i}$ including all activations $a^{(\text{index})}_{\text{component}}$ and blocks $b^{(\text{index})}_{\text{component}}$ performed by the learner during the learning episode. For instance, one might intervene to activate C at 100 ms and B at 1200 ms respectively and observe two subsequent A events. We write this as $\mathbf{d_{demo}}\{d^{(1)}_A = 1500, d^{(2)}_A = 2800|\,\text{Do}[a^{(1)}_C = 100, a^{(2)}_B = 1200]\}$. In order include blocking actions to the intervention set, we indicate their onset and offset times as in $\text{Do}[\{b^{(1)}_A = \{1200, 2400\}, b^{(2)}_A = \{4300, 9000\}]$ meaning that A is blocked from 1200 ms to 2400 ms and again from 4300 ms to 9000 ms.

An immediate issue in calculating the likelihood of the data given a candidate structure is that there are likely to be multiple potential paths of actual causation through which a particular causal structure could have produced the data (Halpern, 2016). Each of these has its own likelihood. For example, if the true structure is a collider $B \to A \leftarrow C$, the data $\mathbf{d_{demo}}$ above might be produced in two ways. C could have caused the first activation of A and B the later one $(Do[a^{(1)}_C] \to d^{(1)}_A, Do[a^{(1)}_B] \to d^{(2)}_A)$. Alternatively, C could have caused the later activation of A, and B the earlier one $(Do[a^{(1)}_C] \to d^{(2)}_A, Do[a^{(1)}_B] \to d^{(1)}_A)$. To construct the total likelihood of a hypothesized causal structure and interventions producing a set of events, we must consider all possible causal paths $\mathbf{Z}_s$ that could describe what actually happened given structure s and then repeat this for every $s \in \mathbf{S}$. Since each path is exclusive and exhaustive conditional on the structure under consideration s, we can sum the path likelihoods to calculate the total likelihood of that structure producing the data:

$$p(\mathbf{d}|s;\mathbf{i}) = \sum_{z' \in \mathbf{Z}_s} p(\mathbf{d}|z';\mathbf{i}). \tag{2}$$

To construct the possible paths, each effect event must be attributed to exactly one preceding event occurring at a component with a causal link to that effect in structure s. Assessing the likelihood of each valid path includes two parts: (1) explaining all actual effects; (2) explaining away any expected effects that did not occur. The first part is just the product of the gamma densities for all the causal delays between observed effects and their putative causes. Each delay is given by

$$\gamma_{pdf}(t_{\text{effect}} - t_{\text{cause}}, \alpha, \beta). \tag{3}$$

For the latter part, we need to check that each cause event assumed by the hypothetical structure has its corresponding effect(s) in the path. Each one that is missing must have failed (with probability $1 - w$), or (with probability w) be either yet to occur or have been blocked from occurring. Combining these possibilities we get the following expression:[2]

$$w\Big[\underbrace{\gamma_{cdf}(\max(0, t_{\text{block}_{\text{onset}}} - t_{\text{cause}}), \min(t_{\text{block}_{\text{offset}}} - t_{\text{cause}}, t_{\text{now}} - t_{\text{cause}}), \alpha, \beta)}_{\text{activation was blocked}} + \tag{4}$$

$$\underbrace{(1 - \gamma_{cdf}(t_{\text{now}} - t_{\text{cause}}, \alpha, \beta)}_{\text{activation has not occurred yet}}\Big] + \underbrace{(1 - w)}_{\text{activation failed}} \tag{5}$$

Thus, the likelihood for a particular causal path given a particular causal structure can be calculated exactly via a combination of diagnostic reasoning — attributing exactly one cause for each observed effect — and predictive reasoning — attributing exactly one effect or failure to each causal link coming out of each activated component.

Cost of inference

There are various ways to measure the computational costs of integrating causal structure evidence. Our inference framework works by considering the various pathways connecting the interventions and effects under each considered structure. The number of paths scales rapidly with the number of plausibly-related effects (Figure 2), meaning a naïve realization of our ideal observer performs an amount of computation that scales super-exponentially in the total number of events observed so far. Nevertheless, considering all past events back to the beginning of time, which we call the *global event* set, is clearly infeasible outside of very simplified toy settings. Inevitably, practical constraints come into play such as excluding from consideration events that occurred long enough ago to have a negligible chance of having caused the most recent effect. For simplicity, in our primary analyses we simply assume learners are focused on a 4-second “backtracking window” (c.f. Gerstenberg et al., 2013). That is, we assume learners enumerate and consider causal pathways involving events or interventions from up to 4 seconds prior to the moment at which the inference is taking place. We chose 4 seconds as the window size as this is long enough to include all plausible causes for any newly occurring event under our delay regime. We refer to these as the *local event* set and assume the learner reasons over a rolling window of local events throughout the trial. This results in a measure of inferential cost

[2] $\gamma_{cdf}(x, y, \alpha, \beta)$ in 5 denotes the cumulative probability of a delay being between x and y in length.

Algorithm: Ideal Observer

```
# Interventions I; Effects D; Events N = I ∪ E; Structures S
Function IdealObserver(N, D, S):
    foreach s ∈ S do
        k = 0
        P = {}
        li[s] = PathwayBuilder(P, N, D, s, k)
    return li
Function PathwayBuilder(P, N, D, s, k):
    if k > |D| then
        return PathwayProb(P)
    else
        C = CauseCandidates(N, D[k], s)
        pr = {}
        foreach c ∈ C do
            P′ = AddingCause(P, D[k], c)
            pr[c] = PathwayBuilder(P′, N, D, s, k + 1)
        return sum(pr)
```

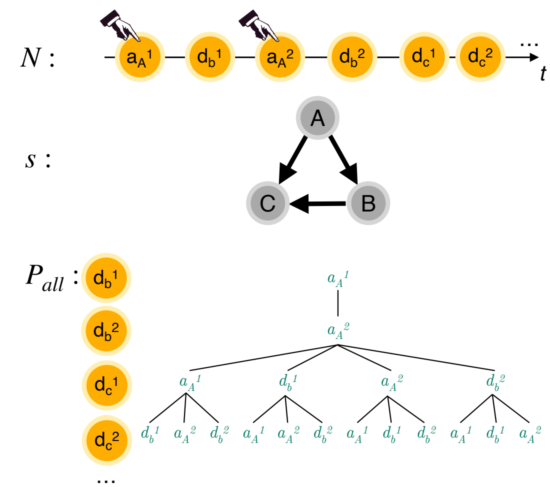

Figure 2

Ideal observer inference algorithm. Ideal Bayesian inference considers each possible structure hypothesis $s \in S$ and every possible causal path $p \in P_s$ that could describe how that structure produced the observations. The number of possible paths grows rapidly in the number of events as illustrated on right hand side with an example recursion tree showing all paths connecting events $a_A^1 ... d_B^2$ conditional on the structure $C \leftarrow A \rightarrow B \rightarrow C$.

that shifts throughout a trial as a function of the number of recent events (see Figure 3). We will also examine other choices of the window size in our later experimental analyses.

While idealized Bayesian inference also requires estimation of the evidence in parallel under all possible hypotheses, in practice it is implausible that a bounded learner would consider the entire hypothesis space at any one time since this quickly becomes intractable as the number of components increases. For instance, there are 4064 possible structures linking 4 components together and 1048576 for 5 components. A recent proposal for how learners mitigate the complexity of structure inference in the natural world is that they consider hypotheses sequentially. For example, in the atemporal dataset setting, it has been argued that participants consider evidence under a single favored hypothesis at a time, regenerating or adapting this hypothesis only to the extent that it fails to explain the most recent evidence (Bonawitz et al., 2014; Bramley, Dayan et al., 2017).

Since humans must, by necessity, find a more scalable approach to causal inference than our normative algorithm in order to to succeed in the wild, we think of the idealized Bayesian inference as an upper bound on computational cost of inference. We thus explore intervention behavior under several plausible inference-complexity-scaling functions based on either the global or local number of events n and some base parameter c. Concretely, we consider linear $O(n)$, polynomial $O(n^c)$, exponential $O(c^n)$ scaling in the events under consideration. These functions differ at how fast the cost grows with the increase of event

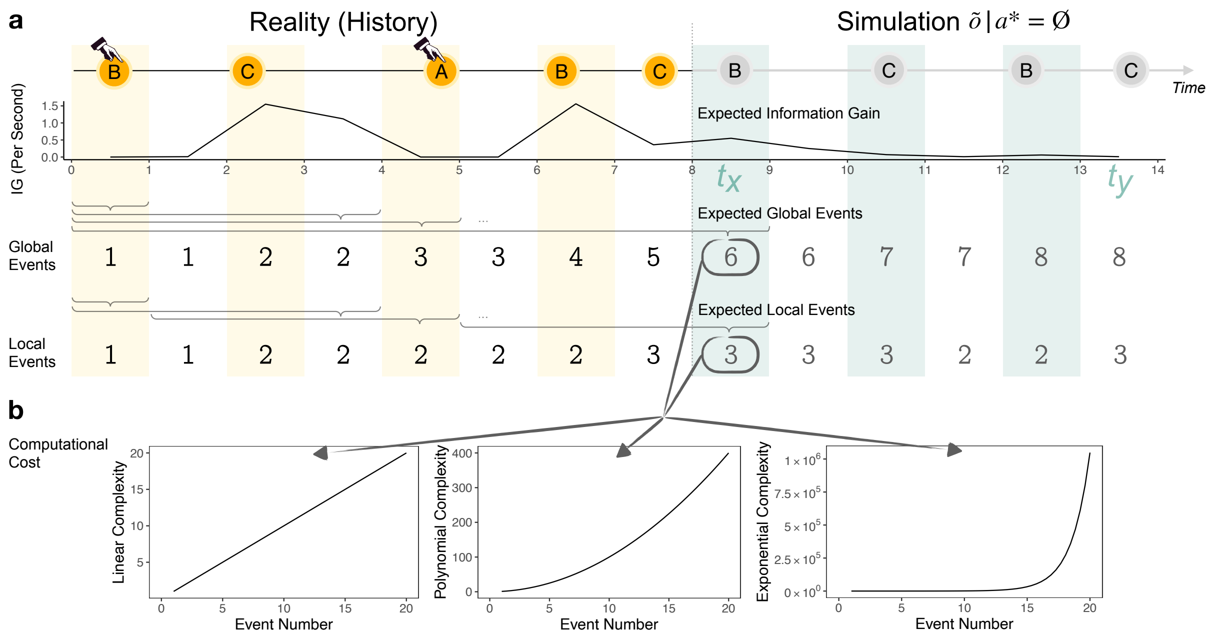

Figure 3

An illustrative example of historical and expected *Information Gain (IG), alongside historical and expected global events and local events. a) Yellow portion (left of t_x) = Frame-by-frame IG about true structure; Global events (since start of observation); and Local events (within the last 4 seconds). Green portion (right of t_x) Expected upcoming information, global and local events. b) Three possible computational cost functions of event number.*

numbers (Figure 3b).

Intervention selection

Intervention selection is the problem of choosing what to do now, in order to support future learning. Normatively, this depends on the prior $P(S)$ at the point of the decision, which in turn depends on the already-observed data $\mathbf{d}_t$ and earlier interventions $\mathbf{i}_t$. In our first experiment where participants can activate but not block, the participant must choose, at each moment in time, between intervening on one of the components or doing nothing: leading to the action space $\mathbf{A} = \{a_A, a_B, a_C, \varnothing\}$ for three node systems. If the learner has used all activation chances, this reduces to just the option of doing nothing $\varnothing$. In our second experiment, where participants and models can also block components the action space is larger including actions that toggle the block status of each node such that it becomes blocked if currently unblocked or unblocked if currently blocked (e.g. $\mathbf{A} = \{a_A, a_B, a_C, b_A, b_B, b_C, \varnothing\}$).[3]

[3] While in principle this decision needs to be made every moment in time, in practice we simplify our analyses by assuming that learners make one intervention selection decision per second. This is an assumption we discuss in our analyses and return to in the General Discussion.

Information gain (IG) is a common currency for measuring the value of evidence for an ideal learner (Coenen, Nelson et al., 2019; Nelson, 2005; Shannon, 1948). The goal is to select the intervention (or sequence of interventions) that is *expected* to have high information gain, or in other words, to best reduce the learner's uncertainty. To do this exactly, one must quantify how much every possible intervention decision i_t^* is expected to reduce future uncertainty about the structure of the causal system given the current beliefs and marginalizing over consideration of possible future evidence. We take a greedy approach in favoring actions expected to maximally reduce future uncertainty at this point but without considering potential subsequent actions.[4] The learner's uncertainty at time t_x can be measured by calculating the Shannon entropy $H(S)_{t_x}$ of the current prior $P(S)_{t_x}$ based on all the evidence experience so far:

$$H(S)_{t_x} = \sum_{s \in S} P(s)_{t_x} \log_2 \frac{1}{P(s)_{t_x}} \tag{6}$$

The ideal calculation of future information should consider all possible future evidence $\mathbf{o}$ up to some future time point t_y, given the hypothetical action i^*. However, unlike the contingency setting, the outcome space here is continuous meaning we must approximate this integral by sampling a subset of possible futures. We achieve this by simulating a set of possible outcome sequences $\tilde{\mathbf{o}}$ under different structures. The number of samples simulated under each structure is based on the structure's prior probability (Nelson, 2005). For each simulated future $\tilde{o} \in \tilde{\mathbf{o}}$ we can compute the information gain as:

$$\mathrm{IG}(i^*, \tilde{o})_{t_y}^{t_x} = H(S)_{t_x} - H(S, i^*, \tilde{o})_{t_y} \tag{7}$$

and expected information gain as:

$$\mathrm{EIG}[i^*]_{t_y}^{t_x} = \sum_{\tilde{o} \in \tilde{\mathbf{o}}} \mathrm{IG}(i^*, \tilde{o})_{t_y}^{t_x}. \tag{8}$$

Note that in this setting, the anticipated information results not only from the focal action choice i^*, but also from other recent actions and effects that may still be expected to produce further effects and evidence. This means that one can often expect substantial information to be forthcoming even when choosing not to act ($i^* = \varnothing$).

Figure 4a shows an example where the learner has already intervened, activating C at $t = 0$. Even though no effects have occurred yet, they are considering what to do one second later, at $t = 1$. The value of doing nothing ($\varnothing$) is relatively low at this point as it only includes expected information resulting from the previous intervention. The learner expects less value from activating C a second time than for activating something else, since they expect to learn about the consequences of C from their first intervention. The blocking actions ($\{b_A, ..., b_C\}$) also have low expected information since, at this stage, they would only serve to block potentially informative dynamics produced by the previous intervention.

[4] This is a common choice due to submodularity results about the diminishing utility of planning ahead in active learning problems (Guillory, 2012).

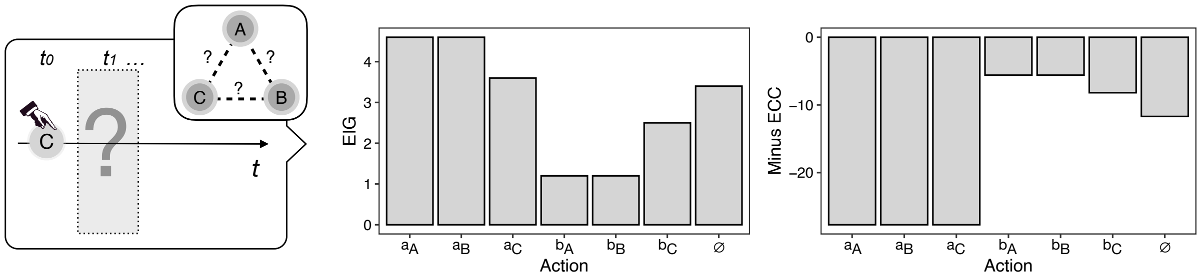

Figure 4

Example of expected information gain (EIG) and expected computational cost (ECC). The learner activated C at t_0 and is now deciding what to do at t_1. The notions of a_X, b_X, and $\varnothing$ stand for choices to activate X, block X, or do nothing, respectively. Both EIG and ECC are temporally discounted. ECC was calculated based on expected global events with a linear function.

Similar to expected information gain, we can also anticipate computational cost of integrating future evidence (CC). This involves counting the events occurring in simulated outcomes $\tilde{o} \in \tilde{\mathbf{o}}$. For each hypothetical future time point considered, we count the recent events $n(t)$ and compute the consequent complexity of performing inference about how these events could relate:

$$\mathrm{CC}(i^*, \tilde{o})_{t_y}^{t_x} = f_{\text{complexity}}(n_{t_y}^{t_x}, c). \tag{9}$$

where we will later allow $f_{\text{complexity}}$ to be of linear, polynomial or exponential form with some parameter c, in either the anticipated local or global events (see Figure 3).

We can then compute *Expected Computational Cost* (ECC) by summing over $\tilde{o} \in \tilde{\mathbf{o}}$

$$\mathrm{ECC}[i^*]_{t_y}^{t_x} = \sum_{\tilde{o} \in \tilde{\mathbf{o}}} \mathrm{CC}(S, i^*, \tilde{o})_{t_y}^{t_x} \tag{10}$$

According to the resource-rational framework (Lieder & Griffiths, 2020), the expected utility of an action $\mathbb{E}[U(i^*)]$ to a bounded learner balances expected reward and cost of computation. In our case, this results in the following equation:

$$\mathbb{E}[U(i^*)_{t_x}] = \sum_{t=t_x}^{t_y-1} R(t) \cdot \left[\mathrm{EIG}[i^*]_t^{t+1} - \omega \cdot \mathrm{ECC}[i^*]_t^{t+1}\right] \tag{11}$$

where we assume a 1 second granularity for measurement, and where ω scales the cost component to align it with the epistemic reward scale of bits, the sum aggregates the expected future gains and costs over future seconds up until t_y, with $R(t)$ as a discount function which diminishes the utility of information and the dis-utility of computational costs the further into the future they occur. In our case this is simply done according to how long the trial remains to end (i.e. chance to affect the bonus):

$$R(t) = 1 - \frac{t}{t_{end}} \tag{12}$$

The ideal t_y horizon should be the end of the learning episode t_{end} (i.e. 45 s in our experiments), but we found no substantial impact upon our choice predictions beyond $t_x + 6$.[5] Finally, a resource rational learner should behave according to:

$$\underset{i^* \in \mathbf{A}}{\operatorname{argmax}} \left[\mathbb{E}[U(i^*)_{t_x}] \right] \tag{13}$$

Figure 3 visualizes the various elements of a trial that combine into our resource rational algorithm and Figure 4 shows an example in which information gain and inferential complexity differ in the choices they favor and hence trade-off in shaping the choice. In sum, our framework captures how a resource rational agent should decide when and where to intervene to support their causal structure learning. We will compare human interventions against the predictions of this modeling framework.

Overview of experiments

We conducted two experiments to test how people actively learn causal structures in continuous time. In both, we manipulated the reliability of the cause-effect delays and included a range of cyclic and acyclic causal structures. In Experiment 1, participants were only able to activate components while in Experiment 2, we additionally allowed learners to block components as well. We hypothesized that performance would be lower in the irregular delay condition given that evidence was formally weaker, and that people have been found to make weaker attributions when putative cause–effect delays have high variance across instances (Greville & Buehner, 2010). We also hypothesized that performance would be worse in cyclic systems given the likely increase in events, interdependence and concomitant complexity. However, we expect this to vary as a function of the quality and reactivity of participant's intervention choices. Thus, we will also examine participants' active learning as a function of the above factors, similarly asking how it differs across structures, delay conditions and assess the extent to which it is reactive in anticipation of trading off information gain and evidential complexity. We will then compare model predictions to the data from both experiments. This will allow us to characterize what strategies people use to decide when and where to intervene.

---

[5] Intuitively, this horizon is reasonable here for several reasons: (1) The rational temporal discount factor makes distant future less important. (2) Expected information gain under the "greedy" assumption of no future activations approaches zero after a handful of seconds, by which time even the most complex causal systems have had enough time to loop through all their causal relationships at least once. (3) The inherently stochastic delays combined with the chaotic causal interactions and compounded by the learner's uncertainty thereof leads to chaotic simulated dynamics whose predictive power rapidly drops toward chance beyond a few seconds (cf. Bramley, Gerstenberg, Tenenbaum et al., 2018).

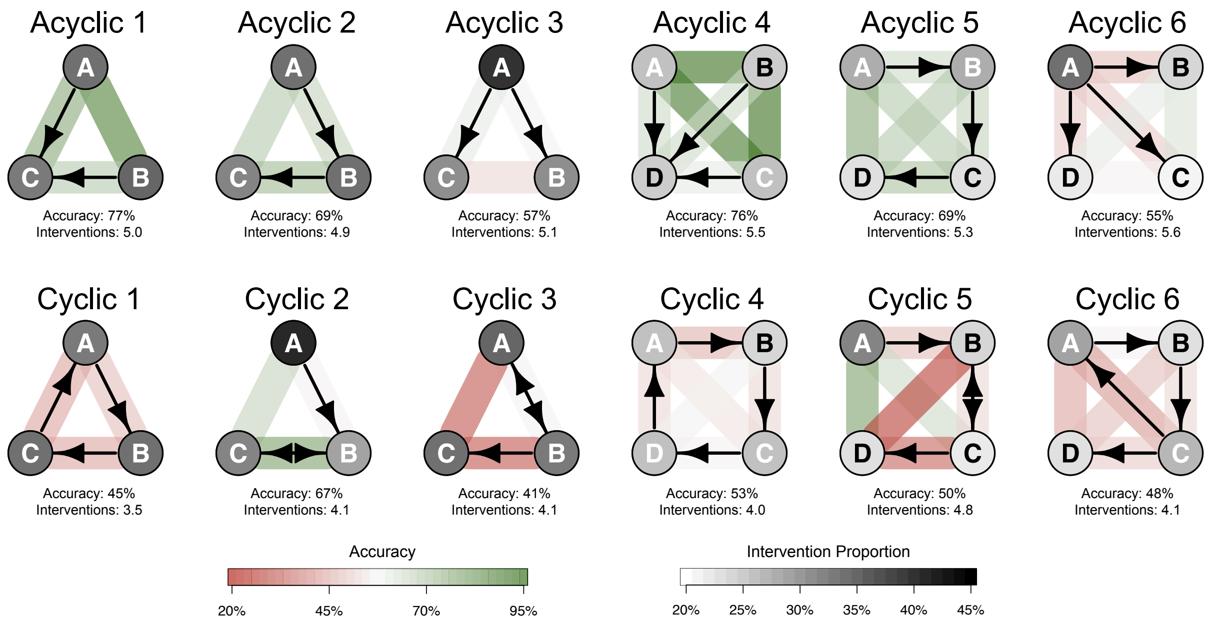

Figure 5

Devices tested and results from Experiment 1. Edge shading indicates accuracy. Node shading indicates intervention choice prevalence by component. For interpretation of colors in this figure, the reader is referred to the online version of this article.

Experiment 1

Participants

Seventy-four participants (40 female, 34 male, aged 30 ± 11) were recruited from Prolific Academic and were randomly assigned to either the reliable-delay ($N = 36$) or unreliable-delay ($N = 38$) condition. Participants received a basic payment of £1 and a bonus depending on performance (see Methods). Nine additional participants were tested but removed from the analysis because they left default answers on all trials ($N = 6$) or had at least one trial in which they performed no interventions at all ($N = 3$). The sample size was chosen to be in line with related work on causal learning (Bramley, Gerstenberg, Mayrhofer et al., 2018; Coenen, Ruggeri et al., 2019).

Methods

Design and Stimuli. Each participant learned about 12 test devices with either 3 or 4 components, including 6 cyclic structures and 6 acyclic structures (Figure 5).[6] The acyclic devices were chosen to exemplify a variety of causal patterns including common effects (`Acyclic1` and `4`), chains (`2` and `5`) common causes (`3` and `6`). The cyclic devices were chosen to match the approximate number of edges in the acyclic systems but to

[6] The task is available here: https://eco.ppls.ed.ac.uk/~s1940738/demo/time_intervention/

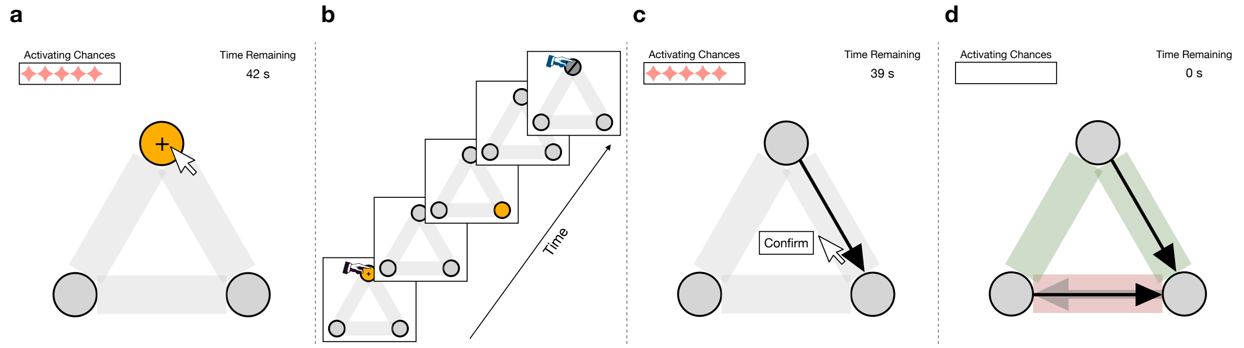

Figure 6

Experimental procedure. a) Experimental interface. Up to 6 interventions could be performed by clicking on the components during the 45 second trial. b) Example timeline. Interventions lead to subsequent activations determined by the direction and delay of the causal connections in the true model. c) Judgments. Participants can indicate their beliefs about the structure during the trials by clicking on the edges. Participants in Experiment 1 need to click the confirm button to lock their answers. d) Feedback. At the end of each trial feedback was provided (green = correct; red = incorrect; wide gray arrows in the background indicate ground truth).

explore some potentially interesting cyclic patterns. These include full loops (`Cyclic1` and `4`) and short loops with incoming (`2` and `5`) and outgoing (`3`, `5` and `6`) edges. There were two between-subject conditions: *reliable* in which the genuine cause-effect delays were sampled from a Gamma with 1.5 ± 0.1s and *unreliable* in which they were 1.5 ± 0.7s (Figure 1d). During each trial, participants could perform up to 6 activation interventions. All causal links worked 90% of the time and no events occurred without being caused by an intervention or other event (i.e. none of the components activated spontaneously by themselves).

Procedure. Prior to the inference task, participants completed instructions, practice trial and comprehension checks. They were not given a particular cover story but simply told they would be investigating the causal structure of a number of abstract "devices". In the instructions, participants were trained on the true cause-effect delays in their condition with text and shown a video example of a device with its causal links revealed. They were then trained on how to register structure judgments. Participants learned that they would receive a £0.03 bonus for each connection correctly marked at a randomly chosen and unmarked point during each trial (for a theoretical maximum total bonus of £1.62). This was emphasized in the instructions to encourage participants to mark connections as quickly as possible. Participants had to correctly answer 5 comprehension check questions before proceeding to the main task. Finally, participants completed a practice trial on a device with a collider structure (`Acyclic1` in Figure 5).

In the main task, participants faced the 12 test devices in random order with randomly orientated and unlabeled components. For each device, participants were

Table 1

Accuracy Separated by Conditions.

	Delay Reliability		Structure Cyclicity			Structure Node	
	Reliable	Unreliable	Unlinked	Acyclic	Cyclic	3-node	4-node
Experiment 1	62%±33%	55%±32%	–	67%±33%	51%±31%	59%±33%	59%±30%
Experiment 2	61%±35%	61%±34%	77%±39%	65%±33%	52%±32%	60%±37%	61%±31%

presented with the 3- or 4-components visualized as gray circles evenly spaced on a white background (Figure 6a). They then had 45 seconds to learn about how the components were connected. During this time, participants could intervene and activate components by left-clicking on them up to 6 times. Intervened-on components were marked by a “+” symbol (Figure 6a and b). All activated components turned yellow for 200 ms and then returned to gray. Initially, all components were inactive, and no connections were marked between them.

Participants were able to indicate their current belief about the causal structure as often as they liked during each learning problem. To do so, participants clicked on the gray area between components to toggle between a causal connection in either direction, or no connection. Each click cycled through the options ($x \to y$, $y \leftarrow x$, $x \leftrightarrow y$, no relationship) in a random order varied between participants. Participants confirmed their choices by clicking a confirm button that appeared in the middle (Figure 6c). Links did not disappear after being confirmed, so participants were still able to update earlier judgments. Whatever participants had marked at the end of 45s was automatically registered as the final judgment for that trial. At the end of the 45 seconds, participants received feedback showing which connections they had marked correctly or incorrectly and then moved on to the next device (Figure 6d).

Results

We report accuracy by delay condition (reliable vs. unreliable), device type (acyclic vs. cyclic), number of components (3 vs. 4) and conditional on intervention patterns. We compare participants' patterns against normative inference and by considering the computational processing costs. We will then focus on participants' intervention choice behavior asking whether we can understand it as driven by a trade-off between maximizing evidence and minimizing computational cost. Our accuracy analyses use linear mixed-effect models (LMMs) fit with R's `lmer4` package including random effect terms for subject ID and twelve causal structure types. For all LMMs we reported standard coefficient estimates β, t values, significance, and 95% confident intervals (CI).[7]

[7] Data and the analysis code are available here: https://github.com/tianweigong/time_and_intervention.

Accuracy. Figure 5 shows how accurate participants' judgments were separated by device. Participants registered their causal judgments 2.45 ± 1.31 times per trial. Final judgments identified the majority of the causal connections correctly ($62\% \pm 34\%$) but with marked variation across and within devices. Participants' final judgments generally improved on the accuracy of initial judgments in the 79% of trials in which the answer was registered more than once (initial accuracy: $58\% \pm 30\%$, $\beta = 0.11$, $t = 2.81$, $p = .005$, $CI = [0.03, 0.19]$). In the following, we focus on participants' final structure judgments.

Participants performed significantly above chance in both delay groups (chance: 25%, reliable: $t(35) = 10.83$, $p < .001$, $Cohen's\ d = 1.80$; unreliable: $t(37) = 11.44$, $p < .001$, $Cohen's\ d = 1.86$, Table 1) and were above chance for all 12 structures taken individually in both the reliable ($ts(35) > 4.15$, $ps < .001$) and the unreliable condition ($ts(37) > 3.49$, $ps < .01$) with the exception of `Cyclic1` (unreliable: $t(37) = 1.94$, $p = .06$, Figure 5).

Accuracy separated by condition is summarized in Table 1. There was a main effect of structure cyclicity ($\beta = 0.50$, $t = 2.93$, $p = .02$, $CI = [0.18, 0.82]$) but no main effects of reliability nor the number of nodes ($ps > .05$). However, there was an interaction between reliability and cyclicity ($\beta = 0.34$, $t = 2.36$, $p = .01$, $CI = [0.06, 0.62]$). Performance was significantly higher in the reliable condition on cyclic structures (56% vs. 45%) but not acyclic structures (69% vs. 66%, Figure 7). There was also an interaction between delay reliability and number of nodes ($\beta = 0.29$, $t = 2.03$, $p = .04$, $CI = [0.01, 0.57]$). That is, performance was slightly higher in the reliable than in unreliable condition for four-node structures (64% vs. 54%) but this was not the case for three-node structures (61% vs. 57%, Figure 7).

Participants were accurate at judging whether the structure was cyclic or acyclic regardless of whether they reported the exact structure, choosing the correct class 82%±38% of time for acyclic structures and 77%±42% of time for cyclic structures. There was no evidence for a difference in the frequency of mistaking cyclic for acyclic vs. acyclic for cyclic ($t(73) = 1.15$, $p = 0.25$).

In sum, participants' performance was dependent on whether the device they were learning about contained a cycle. Instead of a direct effect of delay reliability anticipated by previous studies (Greville & Buehner, 2010), we found an interaction whereby reliable delays appeared to assist inference in acyclic but not in cyclic systems and for systems with four nodes but not with three.

To better understand participants' performance, we now compare participants' responses to those of an ideal observer following the normative procedure we describe earlier in the paper.

Informativeness. We calculated the accuracy of the ideal observer (IO) based on the 45-seconds of evidence generated by each participant on each trial. This acts as a measure of how *informative* the evidence generated by the participants is. As shown in Figure 7, in contrast to human learners, IO was more accurate on reliable than unreliable trials ($\beta = 0.24$, $t = 2.92$, $p = .005$, $CI = [0.08, 0.39]$). It was also more accurate on cyclic

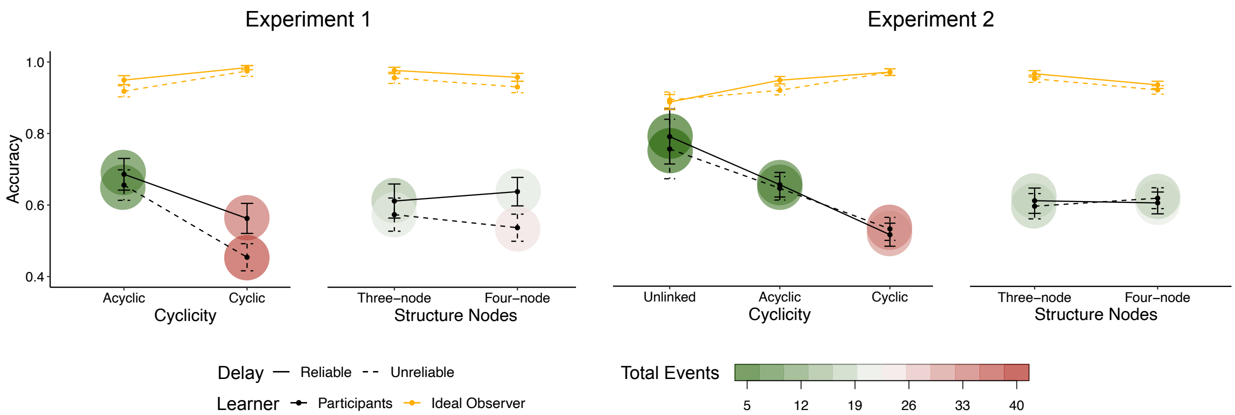

Figure 7

Participants vs. the ideal observer's accuracy and event numbers upon human generated evidence. Error bars indicate 95% confidence intervals.For interpretation of colors in this figure, the reader is referred to the online version of this article.

structures compared to acyclic structures ($\beta = 0.44$, $t = 2.91$, $p = .017$, $CI = [0.16, 0.73]$). There was a three-way interaction ($\beta = 0.59$, $t = 2.31$, $p = .02$, $CI = [0.09, 1.09]$) such that the difference of delay reliability was significantly larger in four-node acyclic structures than others (94% vs. 89%).[8]

IO accuracy therefore does not predict human accuracy (LMM: $\beta = 0.05$, $t = 1.62$, $p = .114$, $CI = [-0.01, 0.11]$). Looking more closely, we find that the relationship between IO accuracy and participant accuracy is moderated by cyclicity ($\beta = 0.16$, $t = 2.80$, $p = .005$, $CI = [0.05, 0.28]$) such that there is a positive association for acyclic structures ($\beta = 0.12$, $t = 3.16$, $p = .002$, $CI = [0.05, 0.20]$) but not cyclic structures ($\beta = -0.04$, $t = 0.89$, $p = .376$, $CI = [-0.14, 0.05]$).

In sum, IO accuracy deviated from human accuracy not only in being substantially higher overall, but also in terms of how the presence of cycles affected performance. Whereas participants performed better in acyclic than cyclic structures, the IO did the opposite. Human judgments were correlated with IO on acyclic learning problems but not for cyclic ones. We hypothesize that this could be because cyclic structures produce more complex evidence thereby creating prohibitive computational demands. In the next section, we consider whether a measure of computational cost can explain the discrepancies between human performance and the evidence available from an IO perspective.

Markers of computational cost. As introduced in the modeling framework above, the computational cost of structure inference scales in the number of events being

[8] For 5% of trials in Experiment 1 and 3% in Experiment 2, participants generated such a high event density that we were not able to calculate the posterior due to there being too many possible causal paths to evaluate ($> 10^{15}$). This tended to happen if a participant intervened very rapidly, particularly on cyclic structures where each event tended to spawn many subsequent events. We omit these trials from current analyses.

reasoned over. The indices in our modeling are meant to vary from moment to moment, while here we need summary measure for each trial to explore whether complexity differences help account for aggregate accuracy patterns. *Total events* occurring in a trial is a intuitive metric of complexity, functionally equivalent here to the average number of events per second. However, given the current online setting where learners need to infer through evidence moment by moment, the distribution of events across the trial is also important. Therefore, we also consider *peak event density*, the moment of highest event density during a trial as an exploratory measure. Intuitively, this measure corresponds to the moment of highest computational load for an online learner attempting to explaining events as they occur. Concretely, for peak event density, we compute mean events-per-second for a rolling 4-second window and take the maximum value for each trial. We show in Table A1 that this measure has a robust inverse-relationship with human accuracy under a variety of choices of window size.

Participants' accuracy was negatively related to total events ($\beta = -0.14$, $t = 3.71$, $p < .001$, $CI = [-0.22, -0.06]$) . Adding manipulated conditions (i.e. delay reliability, structure cyclicity, number of components) and their interactions to the regression model improves its predictive power ($\chi^2(7) = 15.58$, $p = .03$). This means that, the total events can predict accuracy, but it cannot fully account for the deviation between trials. Peak event density is also inversely associated with accuracy ($\beta = -0.24$, $t = 6.29$, $p < .001$, $CI = [-0.32, -0.16]$). For this measure, manipulated conditions (i.e. delay reliability, structure cyclicity, node numbers) and their interactions into this regression model do not significantly improve the model ($\chi^2(7) = 12.44$, $p = .09$). We take this to suggest that moments of high complexity, as indexed by peak event density, are important part of why participants found certain devices much harder to learn than others.

We built an LMM model to predict human accuracy, using subject ID and structure type as fixed effects and including IO accuracy, total events, and peak event density as predictors. Since the intervention count is positively related to IO accuracy, as well as to both total events and peak event density (Table 2), we also included this as a predictor to see whether it could serve as a more direct behavioral index to account for learning quality. Results are shown in Table 3. peak event density significantly predicts human accuracy over and above the information generated (measured by IO accuracy). Importantly, having taken density (as a marker of complexity) into account, IO accuracy also now positively predicts human accuracy. This suggests that successful learning is jointly shaped by both information gain and computational cost.

Intervention choice

We now assess whether participants' intervention choices are qualitatively in line with the predictions of expected information gain and expected computational cost of inference as in our modeling framework. On average, participants performed 4.68 ± 1.46 out of the maximum of 6 interventions on each trial, performing about the same number in

Table 2

Individual Regressions of Using Intervention Count to Predict Relating Trial Measures.

	Experiment 1: Activating			Experiment 2: Activating			Experiment 2: Blocking		
Outcome	β	t	95% CI	β	t	95% CI	β	t	95% CI
IO accuracy	0.23	6.24***	[0.16, 0.31]	0.26	9.87***	[0.20, 0.31]	-0.12	4.80***	[-0.17, -0.07]
Total event	0.34	14.41***	[0.30, 0.39]	0.24	12.42***	[0.20, 0.27]	-0.09	5.08***	[-0.13, -0.06]
Peak local event	0.34	13.53***	[0.29, 0.39]	0.26	13.39***	[0.22, 0.30]	-0.05	2.60**	[-0.08, -0.01]

Note: For activating interventions we used continuous numbers, while for blocking interventions we used binary codes of whether there was blocking behavior (1) or not (0) in each trial due to the non-normally distributed residual. The same for Table 3

Table 3

Linear Mixed-effect Models of Accuracy.

	Experiment 1				Experiment 2			
Predictor	β	t	*sig.*	95% CI	β	t	*sig.*	95% CI
IO accuracy	0.08	2.57	.010	[0.02, 0.14]	0.08	3.62	< .001	[0.04, 0.13]
N Activations	-0.04	1.23	.220	[-0.12, 0.03]	0.10	3.83	< .001	[0.05, 0.15]
> 0 Blocks	–	–	–	–	0.02	0.92	.360	[-0.02, 0.05]
Total events	0.14	1.90	.058	[-0.01, 0.28]	0.004	0.10	.920	[-0.09, 0.10]
Peak event density	-0.37	5.13	< .001	[-0.53, -0.22]	-0.19	3.91	< .001	[-0.29, -0.09]

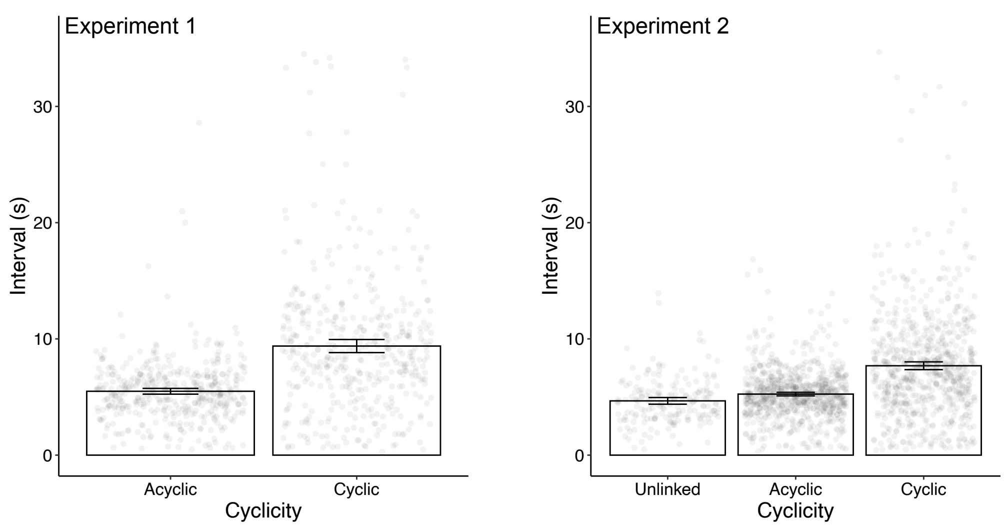

Figure 8

Average time intervals between activating interventions under different structures. Error bars indicate 95% confidence intervals.

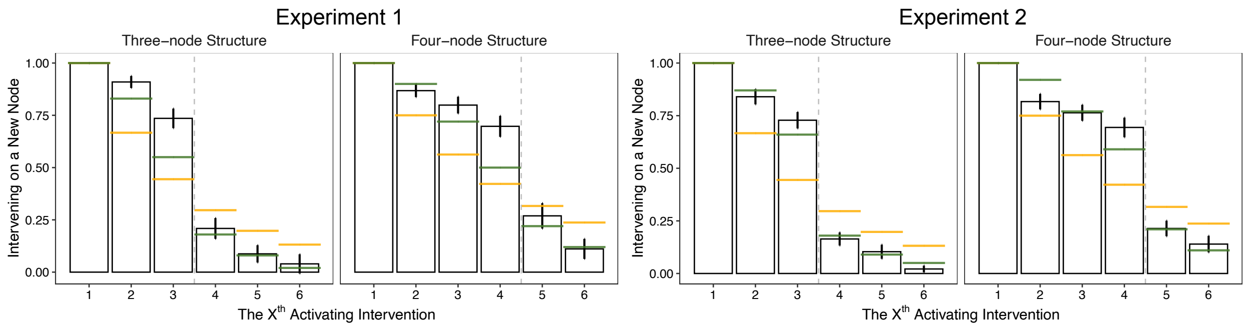

Figure 9

Participants' tendency to activate a node they have not intervened on previously as a function of intervention index. Error bars indicate 95% confidence intervals. Yellow lines indicate the chance under random selection $(N_{node} - 1)^{(X-1)}/(N_{node})^{(X-1)}$ *where* N_{node} *represents the number of nodes in the system. Green lines indicate the level based on the idealized information maximizing intervener who made choices at the same moments as participants and conditional on the same prior evidence. Vertical dashed lines indicates the boundary beyond which the learner has performed enough interventions to have tried every component once. For interpretation of colors in this figure, the reader is referred to the online version of this article.*

the unreliable condition (4.76 ± 1.41) and the reliable condition (4.58 ± 1.51). Participants performed fewer interventions on cyclic problems (4.10 ± 1.50) than on acyclic problems (5.25 ± 1.16, $\beta = 0.79$, $t = 7.22$, $p < .001$, $CI = [0.58, 1.00]$). Participants performed fewer interventions on three-node (4.47 ± 1.54) than four-node problems (4.88 ± 1.35, $\beta = 0.29$, $t = 2.61$, $p = .03$, $CI = [0.58, 1.00]$).

Where to intervene. An efficient sequence of interventions in our unrestricted-hypothesis-space setting intuitively involves both a healthy dose of early exploration — trying each component to learn its effects — but also an exploitative reactive focus — meaning a later tendency to repeat activation of components that appear to have produced effects. This repetition is what allows a learner to gather evidence of variation in the order and delays with which effects propagate through the system, crucial to distinguish between devices with overlapping causal structure. We can look for both of these qualitative features in participants' choices. Figure 5's node shading shows the aggregate proportion of interventions on each node in each structure (recalling that, on-screen, the components were randomly positioned and unlabeled). This shows that, participants' interventions are relatively evenly distributed across components of most devices, but that there is a somewhat higher proportion of activations performed on more causally "central" nodes (i.e. nodes that have many descendant edges; Coenen et al., 2015) in devices that have these such as on A in the two common cause structures (`Acyclic3` and `Acyclic6`).

A marker of early exploration is a tendency to initially sample components to test *without* replacement. That is, choosing something different to activate on one's second test than one's first and so on. To explore whether participants exhibit this pattern, we coded how frequently participants selected a novel component to intervene on, that is, one that they had not yet tested as a function of serial intervention position within the trial. This is shown in Figure 9, and we compared it against chance as well as against the choices of an idealized information maximizing intervener taking actions at the same moments as participants and conditional on the same prior evidence. For both three- and four-node structures, participants were more likely than chance to intervene on untested components until the number of interventions exceeded the number of components in the system ($ts(73) > 10.91, p < .001$). This shows, at a minimum, that participants are not intervening randomly and suggests that they typically begin by exploring the behavior system components that have not activated yet. The informationally efficient intervener shows a similar pattern the first several interventions (Figure 9). Its choices, along with those of participants become reactive to the past evidence in complex ways that do not submit to a straightforward aggregate measure. As such we will examine these choices closely through modeling in a section after the experiments.

When to intervene. As well as choosing where to distribute their up to six intervention choices, participants were also free to choose when to perform them within the 45 second trials. The average interval between each interventions depended on cyclicity ($\beta = 0.79$, $t = 9.48$, $p < .001$, $CI = [0.63, 0.95]$), with participants waiting longer before the next intervention when the structure being learned was cyclic (9.38 ± 5.94s) rather than acyclic (5.49 ± 2.64s, Figure 8). There was no evidence for a difference in this measure between the unreliable and reliable delay conditions (7.55 ± 5.10s vs. 7.26 ± 4.83s) or between three- and four-node problems (7.19 ± 4.92s vs. 7.63 ± 5.01s).

We hypothesized that participants waiting longer under cyclic structures could be motivated by an expectation that future evidence is likely to have a high computational cost, particularly if any additional activations are performed. We investigate this hypothesis at both the structure and the fine-grained action level. On the structure level, we tested whether participants' average waiting time was related to the tendency for a particular structure type to produce complicated evidence (as measured by the typical peak event density produced through interactions with it). Since peak event density measured directly from participants' data depends on the number and spacing of the interventions performed, we instead created an independent measurement of a structure's tendency to produce complex evidence. To do this, we simulated one random intervention for each structure many times and averaging the resultant peak event density. As shown in Figure 10, this reveals that participants' waiting time is indeed related to a structure's tendency to produce complex evidence, with participants waiting 2.58 seconds longer for each additional "typical peak event" ($r = 0.90$, $t(10) = 6.72$, $p < .001$).

We also tested whether participants' moment-by-moment intervention choices are correlated with the anticipated local event count of the near future. To test this, we calculated how many events generated by previous interventions are expected to occur from t_x to t_y (where $t_y = t_x + 6$, see *Intervention Selection*). This provides a baseline cost which fluctuates depending on how much activity the system currently has ongoing and the

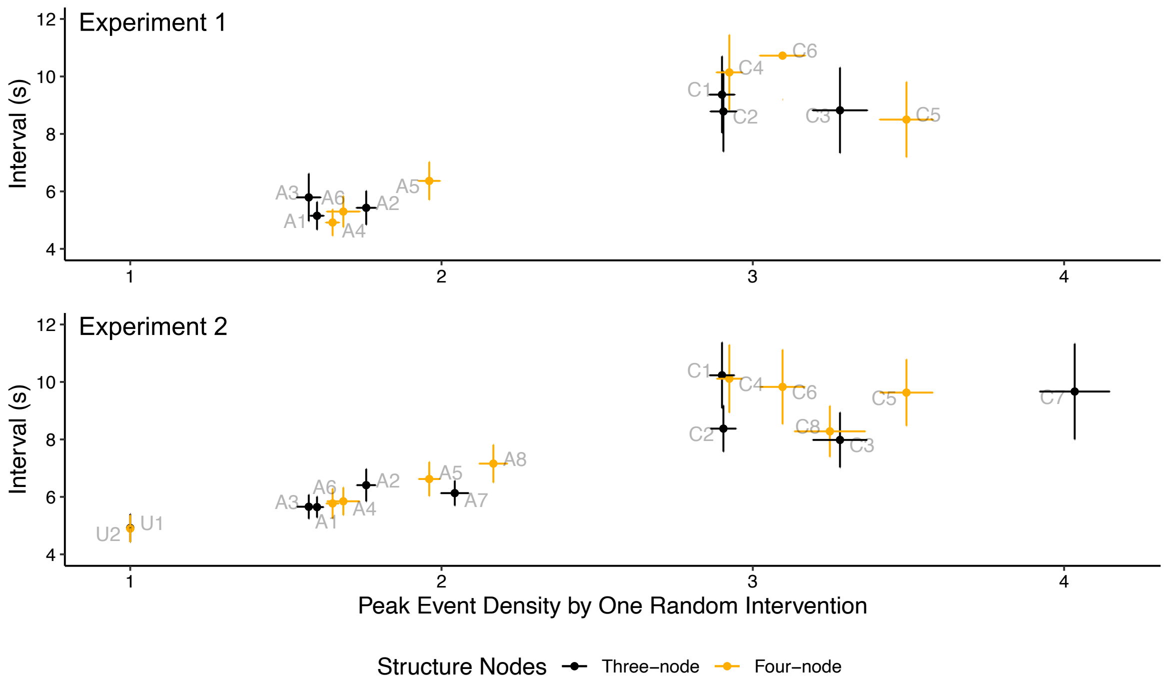

Figure 10

Participants' average waiting time between two activating interventions under different structures. Anticipated Local Events (x-axis) were simulated by randomly activating a component 1000 times under both reliable and unreliable delays. Labels: A = Acyclic, C = Cyclic, U = Unlinked. Error bars indicate 95% confidence intervals.

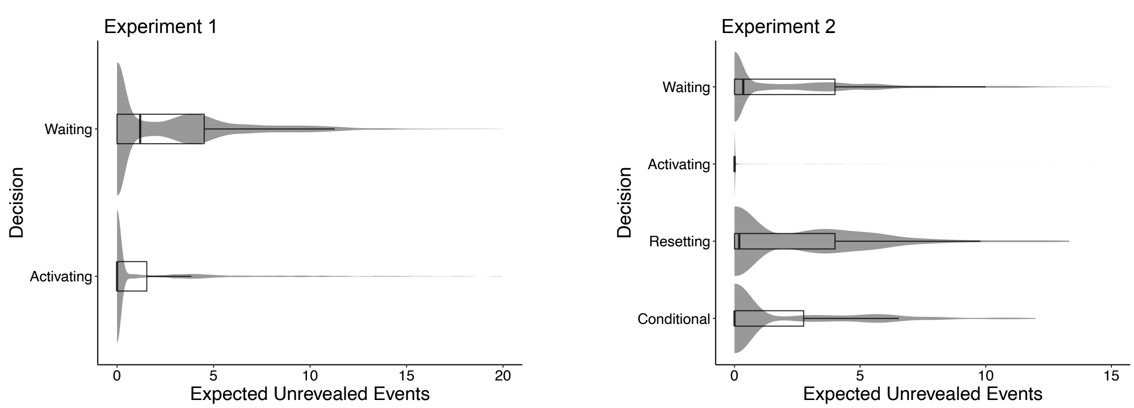

Figure 11

Number of expected unrevealed events across all 1-second decision windows in all trials, as a function of whether an intervention is performed in that window. Windows for which participants performed more than one intervention were excluded. Densities scaled to have equal maximum width. 0.6% and 1.3% of the data, from Experiment 1 and 2 respectively fall outside of the visualized area, i.e. have expected unrevealed events larger than 20.

learner's current prior over structures. We can compare the seconds in which participants do nothing against those in which participants perform an intervention. As shown in Figure 11, people waited — i.e. did not perform any intervention — for 88% of the 1-second windows in which they could have acted (i.e. had not run out of activations yet). Yet in the 12% of time windows where they *did* make activations, the number of expected events ($Median = 0$, $Mean = 1.86$) was lower than those where they did nothing ($Median = 1.25$, $Mean = 2.96$, Mood's median test: $\chi^2(1) = 854.35$, $p < 0.001$). Thus, this is a qualitative marker that participants tended to wait until there was not too much activity expected before intervening.

Discussion

In Experiment 1, we showed that adults are able to infer the causal structure through active intervention in a challenging continuous-time learning setting. We found that participants, unlike the ideal observer model, made more accurate judgments about acyclic structures than cyclic structures. We found that differences in accuracy across conditions could be explained by differences in the character of the evidence, with cases where the interaction resulted in a large spike in the number of local events being associated with lower judgment accuracy. The quality of evidence from an Ideal Observer perspective was only predicted human performance when the influential complexity were taken into account. We take this to support our central idea that managing computational cost plays an important role in interventional decisions in a real-time causal learning setting.

People's when-to-intervene also sought to manage and control computational demands. They waited longer between two interventions in structures that could generate denser events in theory, and choose to wait, rather than activate components to produce more events, when the expected future inferential complexity is already high.

For where-to-intervene, we found evidence that participants used their interventions to systematically explore the devices. Participants showed a tendency to intervene on causally "central" components once discovered. Note that the role of a root component activation differs in this setting to the atemporal settings studied in the past literature. Interventions on known-to-be causally central components has been framed as a Positive Testing Strategy on the grounds that it is only sporadically informative in the atemporal setting (Coenen et al., 2015; Steyvers et al., 2003). This is because in these settings, multiple causal influences from the root component overshadow one another since all effects are revealed at once, meaning that such positive tests can be among the least informative choices. However, in the continuous time case, intervening on a suspected root may generate rich diagnostic evidence through the delays and order variability (Figure 7; Bramley, Gerstenberg, Mayrhofer et al., 2018). As such, we found that participants' selection of novel components to test was qualitatively in line with the behavior of an efficient information maximizing agent. The extent to which participants' specific where-to-intervene choices reflect an information gain norm will be tested by our model fitting to follow.

Experiment 2

Experiment 2 aims to replicate and extend the results of Experiment 1. This time, participants were not only able to activate components but they could also choose to block components, temporarily preventing them from activating until unblocked again. Intuitively, blocking permits the learner a greater degree of control over interactions with and observations of the system, as they can now isolate components to focus on, and also take control of ongoing activity in the system. On the other hand, a larger action space increases the complexity of the intervention decision making problem and the control required to execute could turn out to be prohibitive. We will examine whether the relationship between event count, ideal and human accuracy is similar to the activation-only setting, and explore how participants use the blocking function. In particular, we will assess whether participants spontaneously use blocks to control the computational cost of inference via reducing the complexity of evidence without substantially reducing its diagnosticity about the causal relationships.

Participants

95 participants (54 female, 40 male, 1 unrevealed, aged 36 ± 12) were recruited from Prolific Academic and were randomly assigned to reliable-delay ($N = 48$) or unreliable-delay ($N = 47$) condition. They received a basic payment of £1 and a bonus depending on performance as in Experiment 1. Fourteen additional participants were tested but removed from the analysis because they answered for all the trials that the structure was completely unconnected, which was the initial default ($N = 7$), or did not perform any interventions in at least one trial ($N = 7$).

Methods

The interface was similar to Experiment 1 with a few changes. In addition to activating components, participants were also able to block components by right clicking on them and to unblock them again with an additional right click. Blocked components were marked visually by turning gray and by showing a stop sign on them (Figure 6b). Blocked components did not activate when they otherwise would have been caused to do so by another event or by a left click activation intervention. Blocked components did not pass on any activations to other components. While participants were limited to 6 activations (identical to Experiment 1), we did not limit how many times components could be blocked and unblocked.

We also extended the test set of causal devices. We added two densely connected acyclic structures (`Acyclic7` and `Acyclic8`, Figure 12), two densely connected cyclic structures (`Cyclic7` and `Cyclic8`), as well as two devices with no causal connection between the components (i.e. `Unlinked1` and `Unlinked2`). The inclusion of unconnected structures served to explore how people intervene under one extreme setting where no effects are ever experienced. While unconnected devices are technically acyclic, they are also qualitatively unique and as such, we treated them as a separate device type in our analyses. The new densely connected structures, by comparison, might produce particularly complex evidence and would so be particularly amenable to the use of blocks. Given the larger set of test stimuli, we did not not include a practice trial in Experiment 2. In other respects, the instructions, incentive structure, and randomization procedure were identical to that of Experiment 1.

We additionally improved the interface in two ways. First, we were concerned that occasionally two activations of a component would overlap making them hard to distinguish visually. In Experiment 1, each activation caused the component to turn yellow for 200 ms, if two activations overlapped this would result in the component appearing yellow for longer but without a clear declination between events. In Experiment 2, we had each component turn yellow and then fade back to gray over 200 ms, this made it easier to detect distinct activation events even if their onset times were very close together. Second, to lower the effort and motor requirements of making judgments, participants were not required to click a "confirm" button to register when they had finished making a change to their structure judgment as they had had to in Experiment 1. One second after they stopped clicking on the edges, the state of their currently-marked structure was automatically registered as their latest judgment.

Results

As in Experiment 1, we first focus on judgment accuracy and then on intervention strategies. We first focus on use of activations and then explore when and where participants use the novel blocking function in this experiment.

Accuracy. Participants registered judgments 4.39 ± 2.43 times per trial. Within trials for which the answer was registered more than once (86% of all trials), final judgments were more accurate than initial judgments with $60\% \pm 34\%$ compared to $48\% \pm 24\%$ of connections correctly identified, $\beta = 0.41$, $t = 14.09$, $p < .001$, $CI = [0.35, 0.47]$. Participants' judgments became more accurate as they approached the end of the trial ($\beta = 0.11$, $t = 10.92$, $p < .001$, $CI = [0.09, 0.12]$). As in Experiment 1, we focus on the final answers as our primary measure of task performance.

Performance in both reliability conditions was significantly above chance (random: 25%, reliable: $t(47) = 11.58$, $p < .001$, Cohen's $d = 1.67$; unreliable: $t(46) = 13.63$, $p < .001$, Cohen's $d = 1.99$, Table 1). The average accuracy for all 18 structures were above chance in the reliable condition ($ts(47) > 4.01$, $ps < .001$) and the unreliable condition ($ts(46) > 3.39$, $ps < .01$) with the exception of `Cyclic7` (reliable: $t(47) = 2.01$, $p = .05$; unreliable: $t(46) = 0.56$, $p = .58$, Figure 12).

Accuracy separated by conditions were summarized in Table 1. Unlike Experiment 1, participants' performance only differed between unlinked and cyclic

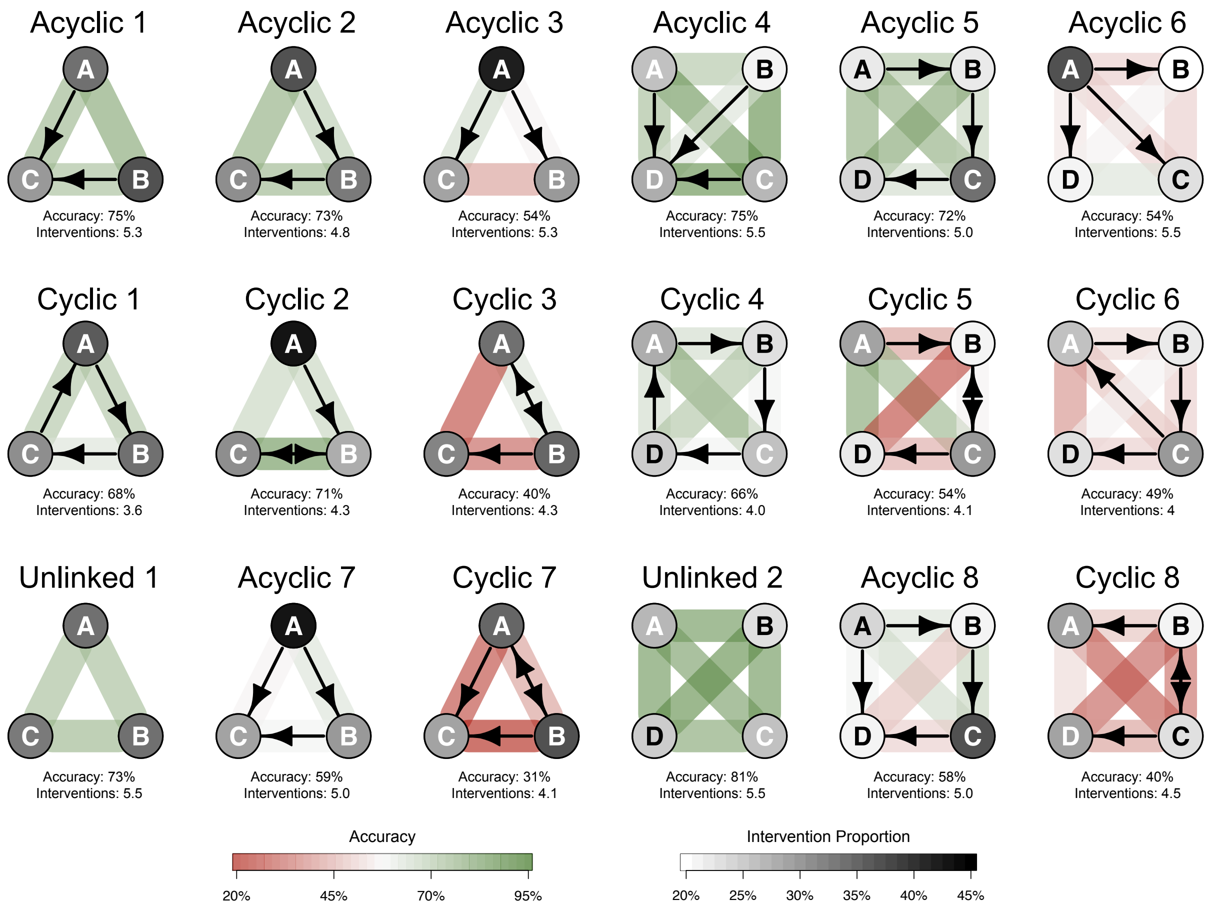

Figure 12

Causal devices results in Experiment 2. Edge shading indicates accuracy. Node shading indicates activating intervention choice prevalence by component. For interpretation of colors in this figure, the reader is referred to the online version of this article.

structures ($\beta = 0.72$, $t = 2.49$, $p = .03$, $CI = [0.10, 1.35]$, Figure 7), with a marginally significant difference between acyclic and cyclic structures ($\beta = 0.37$, $t = 2.01$, $p = .06$, $CI = [0.03, 0.71]$).

As with Experiment 1, participants could generally tell whether the structure was cyclic or acyclic regardless of whether they got the exact structure. They choose the correct class 70%±46% of time for acyclic structures (excluding the unlinked structures) and 80%±40% of time for cyclic structures. Participants more often mistook acyclic structures for cyclic than the reverse ($t(94) = 2.46$, $p = 0.02$).

Informativeness and metrics of complexity. As in Experiment 1, the ideal observer was generally able to make more accurate judgments from evidence involving greater numbers of events. Based on observing participants' interactions, it performed better at identifying the structure of cyclic ($97\% \pm 9\%$) than acyclic ($93\% \pm 12\%$) devices ($\beta = 0.32$, $t = 2.44$, $p = .03$, $CI = [0.08, 0.55]$) and better cyclic than unlinked

(89% $\pm$ 11%) devices ($\beta = 0.73$, $t = 3.57$, $p = .003$, $CI = [0.35, 1.11]$, Figure 7). It also performed better for three-node (96% $\pm$ 10%) than four-node (93% $\pm$ 11%) devices ($\beta = 0.29$, $t = 2.39$, $p = .03$, $CI = [0.07, 0.52]$, Figure 7). Unlike Experiment 1, IO accuracy positively predicted participant judgment accuracy overall ($\beta = 0.07$, $t = 3.11$, $p = .002$, $CI = [0.02, 0.11]$). Similarly to Experiment 1, the degree of correlation differed depending on cyclicity ($\beta = 0.12$, $t = 2.79$, $p = .005$, $CI = [0.04, 0.20]$), with a positive relationship in acyclic devices ($\beta = 0.09$, $t = 2.89$, $p = .004$, $CI = [0.03, 0.15]$) but no relationship for cyclic devices ($\beta = 0.02$, $t = 0.71$, $p = .48$, $CI = [-0.04, 0.09]$).

Total events has an inverse relationship with accuracy ($\beta = -0.08$, $t = 3.10$, $p = .002$, $CI = [-0.14, -0.03]$) as in Experiment 1. Peak event density was also inversely associated with judgment accuracy ($\beta = -0.12$, $t = 4.20$, $p < .001$, $CI = [-0.17, -0.06]$).

Interventions

Participants performed 4.75 ± 1.46 activations on average on each trial. This did not differ across reliable (4.75 ± 1.42) and unreliable (4.74 ± 1.49) conditions or across three-node (4.70 ± 1.51) and four-node (4.79 ± 1.40) structures. However, participants performed fewer activations on cyclic (4.12 ± 1.52) compared to acyclic devices (5.18 ± 1.23, $\beta = 0.72$, $t = 8.43$, $p < .001$, $CI = [0.56, 0.88]$), or compared to unlinked devices (5.53 ± 0.96, $\beta = 0.97$, $t = 7.10$, $p < .001$, $CI = [0.71, 1.22]$). Participants performed blocks on 27% of trials (949 times in 1710 trials, 0.55 ± 1.17 per trial). 75 of 95 participants (79%) used blocking at least once. Given that the frequency of blocking was much sparser than activations we code trials as 1 (used blocks) or 0 (no blocks) in our statistical analyses.

As with Experiment 1, the number of activating interventions was positively associated with IO accuracy, but also with the number of global events, and local events (Table 2). As we might expect, blocking actions were negatively associated with all three measures, reducing the number of events experienced but also the total information available for an unbounded learner. We used an LMM to relate human judgment accuracy to the above variables. Results are shown in Table 3. This reveals that, as with Experiment 1, human accuracy was positively associated with IO accuracy and negatively with peak event density. There was additionally, a positive association between judgment accuracy and the number of activations performed but no relationship with use of blocking or global effects. Unpacking this, we can see that participants learned more successfully in trials where they used more activations and generated more evidence as long as this did not result in a high peak of local events. While blocking behavior was not associated with accuracy directly, this is not entirely surprising if we assume that participants used this functionality adaptively — for instance, incorporating blocks more often in interactions with more complex harder-to-learn devices, and doing so particularly when experiencing costly complex dynamics. We test this conjecture below.

Where to activate. Similar to Experiment 1, participants tended to explore the device initially by activating each component in turn. Participants were again more likely than the chance level to activate an untested component with their second and third activating interventions (and their fourth for four node devices) ($ps < .001$, Figure 9).

When to activate. We tested the intervals between interventions as we did in Experiment 1. Participants waited longer in cyclic than acyclic structures (9.26 ± 5.57s vs. 6.15 ± 2.54s, $\beta = 0.70$, $t = 8.79$, $p < .001$, $CI = [0.55, 0.85]$) or unlinked structures (4.91 ± 2.33s, $\beta = 0.98$, $t = 7.81$, $p < .001$, $CI = [0.75, 1.22]$), and longer in acyclic than unlinked structure ($\beta = 0.28$, $t = 2.23$, $p = .04$, $CI = [0.05, 0.51]$, Figure 8).

We repeated our structure and action level analysis from Experiment 1 to test whether participants were still more likely to wait rather than perform another activation when they expected that processing evidence in the near future would likely have a high computational cost. Participants' waiting time was again correlated with expected future events at the structure level ($r = 0.91$, $t(16) = 8.63$, $p < .001$, Figure 10). There were, again, more expected unrevealed events on seconds where people waited ($Median = 0.44$, $Mean = 2.58$) than those where they chose to activate a component ($Median = 0$, $Mean = 1.29$, Mood's median test: $\chi^2(1) = 2423.60$, $p < 0.001$, Figure 11).

Where and when to block. We analyze whether participants used blocks and, if so, what motivates their use. We fit a logistic mixed effects model predicting the probability of a participant using blocks as a function of trial type, delay reliability, number of nodes. We found that the propensity to block differed neither between the reliable (28%) and unreliable (26%) delay conditions, nor between three-node (29%) and four-node (25%) devices. However, participants were more likely to use blocks in cyclic (33%) than acyclic (21%) structures ($\beta = 0.86$, $t = 4.83$, $p < .001$, $CI = [0.49, 1.23]$). Surprisingly, the propensity to use blocks when facing unlinked (26%) structures did not differ significantly from cyclic or acyclic structures. We had anticipated participants would be less likely to block in unlinked structures since a key function of blocks in this setting is to manage evidential complexity, and in the unlinked structures this is always minimal. We speculated that some uses of blocks might be spurious since we did not limit their use. For example, sometimes participants may have blocked and unblocked components simply to kill time until the end of the trial, especially after they have used up the activating chances and the system has been silent for a while. To further explore how participants used blocking, we focused on categorizing blocking actions, focusing on two plausible goals of blocking that have distinct empiricial signatures: (1) Blocking to reset the device and (2) Blocking in combination with activating to control for confounding causal paths.

For both of these uses, we derived simple operationalizations. We take "Resetting" blocks to those where the learner blocks and then unblocks a component before performing another activation, without activating any other component in the interval while the component is blocked. In the current setting, Resetting blocks serve to short-circuit ongoing chains of causal effects of previous interventions, essentially resetting the mechanism so that subsequent tests can be performed without interference. In contrast, the "Controlling" category includes blocks that appear to be used as a way to perform a controlled test, essentially isolating a sub-network — made up of all the components except

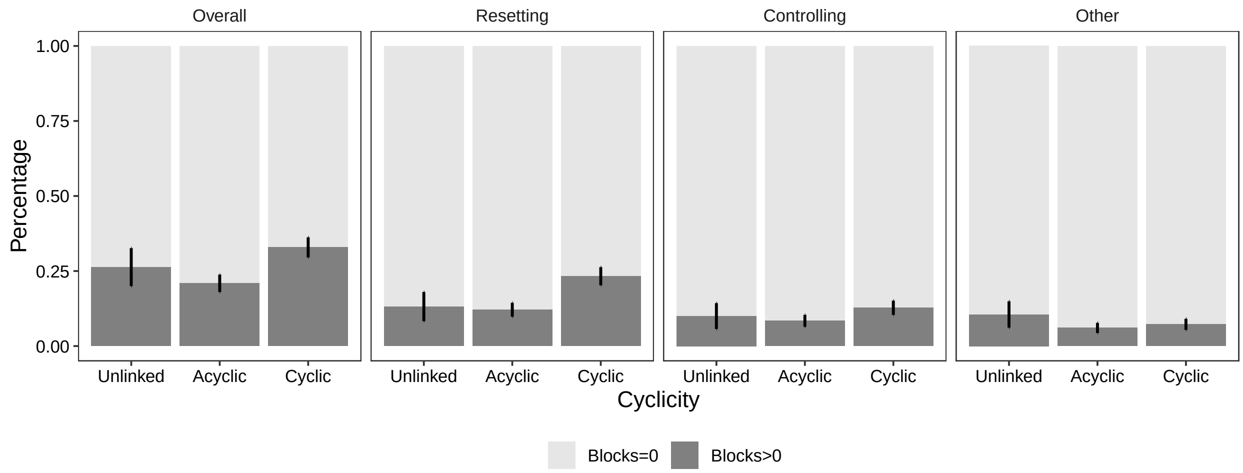

Figure 13

Percentages of blocking behaviors. Error bars indicate 95% confidence intervals.

the blocked one(s) — allowing this sub-network to be investigated through an activation without the possibility of interference from any pathways through the blocked component. Both forms of blocking reduce the local and global events experienced during the trial but do so in conceptually different ways. As such, we categorize each blocking behavior as either Resetting, Controlling, or Other. Resetting blocks — where the next action is to unblock the same component — made up 55%, Controlling blocks —- those followed by an activation of a different component — made up 24% —- and the remaining 21% were classified as Other. This nuisance category includes cases where a block is performed by a participant who has no activations remaining or performs no subsequent activation or unblocking action before the end of the trial.

Participants performed more Resetting blocks in cyclic (23%) than the acyclic (12%, $\beta = 1.03$, $t = 5.66$, $p < .001$, $CI = [0.67, 1.39]$) or unlinked (13%, $\beta = 0.91$, $t = 3.10$, $p = .002$, $CI = [0.33, 1.48]$) devices. There was no difference between acyclic and unlinked structures. Similarly, participants performed more Controlling blocks in cyclic (13%) than acyclic (8%, $\beta = 0.54$, $t = 2.93$, $p = .003$, $CI = [0.18, 0.90]$) devices, but unlinked structures were not significantly different than either (10%). Participants performed slightly more Other-type blocks in unlinked structures (11%) than acyclic structures (6%,$\beta = 0.74$, $t = 2.36$, $p = .018$, $CI = [0.13, 1.36]$, which is marginally significant under Bonferroni correction) but not cyclic structures (7%). In sum, both Resetting and Controlling blocking were used more on cyclic devices.

We further tested whether the results above could be explained by anticipation of complexity. Similar to intervention intervals, we examined the correlation between blocking behaviors and characteristic peak event density for the 18 structures (Figure 14). The proportion of participants using blocking was positively associated with characteristic peak event density overall ($r = 0.64$, $t(16) = 3.32$, $p = .004$), and also for Resetting ($r = 0.74$, $t(16) = 4.39$, $p < .001$), Controlling ($r = 0.63$, $t(16) = 3.27$, $p < .005$) blocks, but not Other

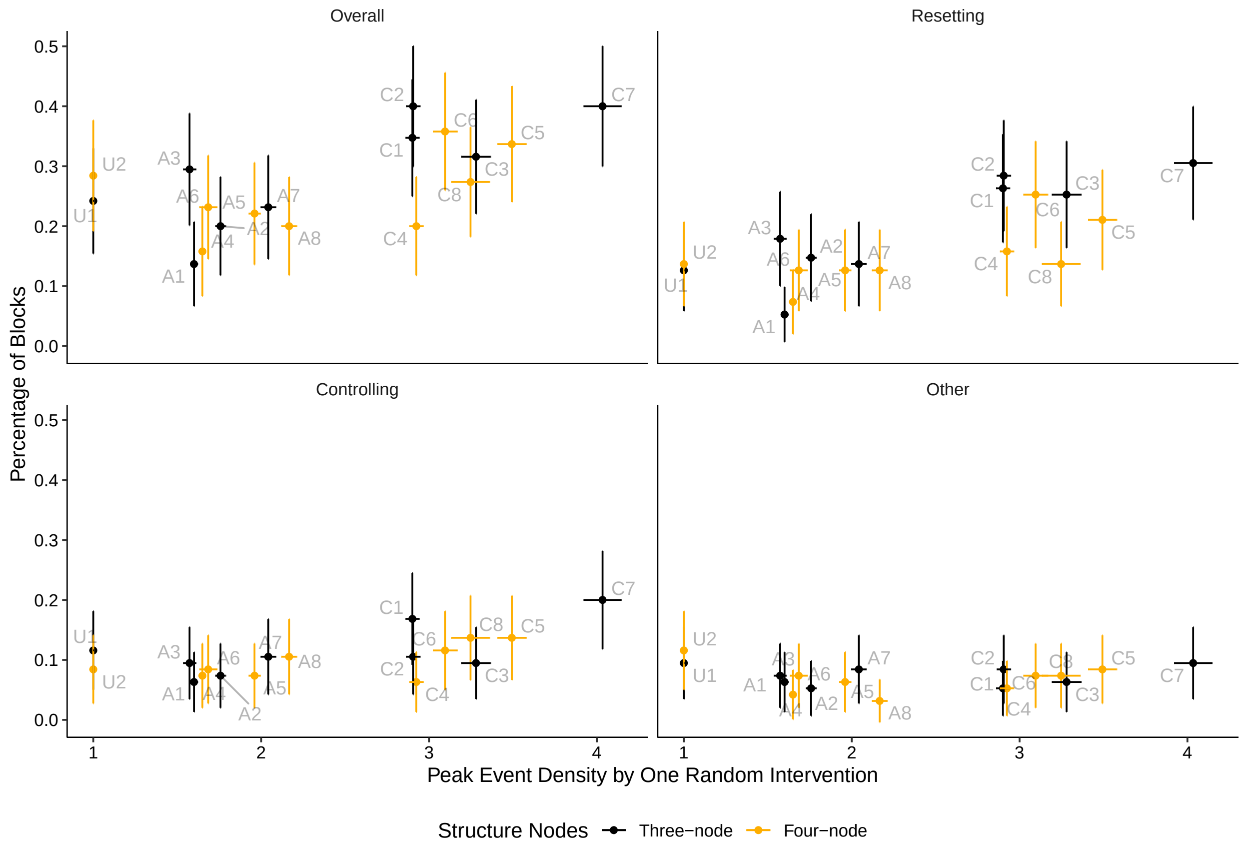

Figure 14

Participants' blocking behaviors under different structures. Data of x-axis was simulated by 1000 times for both reliable and unreliable delay. Labels: A = Acyclic, C = Cyclic, U = Unlinked. Error bars indicate 95% confidence intervals.

blocks ($r = -0.05$, $t(16) = 0.18$, $p = .86$).

Finally, we checked whether participants used Resetting and Controlling blocks in ways that make sense from a bounded rationality perspective. We assume that Resetting is useful at moments where expected future complexity is high, while Controlling blocks should be used when people expected low complexity (i.e. when they were ready for the next activating intervention). In line with this, the number of expected unrevealed events was higher for seconds in which Resetting blocks were performed ($Median = 0.21$, $Mean = 2.62$) than those where Controlling blocks were performed ($Median = 0$, $Mean = 2.12$, $\chi^2(1) = 14.69$, $p < 0.001$, Figure 11).

Discussion. In Experiment 2, we allowed participants to use blocking as a tool to facilitate their causal learning. The addition of blocking made the action space larger but also gave participants more fine-grained control over the learning input, allowing them to not just inject excitation into the causal system but also to inhibit it. We replicated the finding from Experiment 1 that higher event density is associated with lower accuracy, but found that, this time, accuracy was less strongly associated with delay, cyclicity, and number of components. This may in part be due to the fact that blocking helped participants to accommodate and counteract the differences in excitability and ambiguity characteristic of interactions with the different causal devices. We also replicated our findings that people waited longer to perform their next activation in structures where expected computational cost was generally higher and that they tended to explore the components systematically with their initial interventions.

Participants used blocking in a quarter of trials but we found that when blocking was employed it was used in sensible ways that primarily managed inferential complexity. They blocked more often in cyclic than acyclic devices and did so when many events could be expected to occur in the near future. This is consistent with the assumption that learners take minimization of computational cost into consideration when choosing how to intervene in real time. We categorized blocks according two potential goals: Those that reset the system and those that combine with an activation to test a subpart of the system. Both of these were associated with moments of high expected complexity, while other undefined blocks were not.

Model Fitting

As a final close analysis of participants intervention selections, we compare a number of models to participants' sequences of choices. All our models assign a probability to all possible actions across a series of discrete windows throughout each trial. Concretely, each model produces a probability distribution over potential actions at each of 45 1-second windows in each trial.[9] We use the action space defined in *Intervention Selection*: $\mathbf{A} = \{a_A, a_B, a_C, (a_D), \varnothing\}$ (Experiment 1) or $\mathbf{A} = \{a_A, a_B, a_C, (a_D), b_A, b_B, b_C, (b_D), \varnothing\}$ (Experiment 2) where a_X denotes activations of component X and b_X represents toggling of Xs blocked status and the bracketed actions are available only for four node problems. As in Experiment 2, all components are taken to be initially unblocked.

We tested cost-free models that only consider EIG, and cost-dependent models that additionally consider the cost of inference as a function of the global or local events combined with one of three complexity functions: linear, polynomial, and exponential. The complexity function controls how quickly the model expects computational costs to increase as the event number increases. For intervention decisions, this affects how sensitive the model predicts people will be in favoring choices less likely to result in large numbers of events in the future. For example, a exponential-cost agent will become highly conservative beyond a certain degree of anticipated complexity will strongly prefer to do nothing or block rather than add any more activity to the system. In contrast a linear function

[9] Thus, we do not attempt to predict when, within a specific 1-second window, any action would be taken but just what action, if any, is performed in each window. Occasionally participants performed more than one action in a 1-second window. This was very rare though, occurring in only 0.41% of windows in Experiment 1 and 0.36% in Experiment 2. For simplicity, we simply treated these multi-action windows as missing data and modeled the other >99% of trials.

Table 4

Intervention Model Fits.

	CV	BIC	τ	ω	θ	N (CV/BIC)
Experiment 1						
Cost free						
Baseline	-17033	34027	0.31			4/14
EIG	-16588	33109	3.01		10.80	5/7
Global cost						
EIG-ECC$_{\text{Linear}}$	-16550	33040	2.00	1.00×10^{-1}	6.80	1/1
EIG-ECC$_{\text{Polynomial}}$	-16456	32880	2.67	4.89×10^{-3}	8.99	9/3
EIG-ECC$_{\text{Exponential}}$	-16561	33062	2.81	1.27×10^{-10}	10.05	2/0
Local cost						
EIG-ECC$_{\text{Linear}}$	-16524	32994	1.82	1.32×10^{-1}	6.13	8/1
EIG-ECC$_{\text{Polynomial}}$	**-16415**	**32792**	2.27	8.82×10^{-3}	7.91	43/48
EIG-ECC$_{\text{Exponential}}$	-16564	33076	2.77	1.25×10^{-10}	9.98	2/0
Experiment 2						
	CV	BIC	τ	ω	θ	N (CV/BIC)
Cost free						
Baseline	-43042	85997	0.27			2/3
EIG	-40507	80887	1.98		8.13	9/51
Global cost						
EIG-ECC$_{\text{Linear}}$	-40507	80898	1.98	3.33×10^{-6}	8.13	0/0
EIG-ECC$_{\text{Polynomial}}$	-40507	80898	1.98	1.69×10^{-6}	8.13	8/2
EIG-ECC$_{\text{Exponential}}$	-40508	80898	1.98	4.94×10^{-14}	8.13	3/0
Local cost						
EIG-ECC$_{\text{Linear}}$	-40507	80898	1.98	3.76×10^{-6}	8.13	6/0
EIG-ECC$_{\text{Polynomial}}$	**-40462**	**80804**	1.86	1.59×10^{-3}	7.63	52/39
EIG-ECC$_{\text{Exponential}}$	-40506	80893	1.95	1.81×10^{-11}	8.03	15/0

Note: BIC for the fully random baseline was 97210 in Experiment 1 and 262515 in Experiment 2. Parameters reported were based on BIC results.

anticipates cost to rise more smoothly with events and so is less dramatic in how it trades off between anticipated complexity and evidence. Global and local costs differ in terms of being based in the global or local events (Figure 3). The global and local models make similar predictions under a linear complexity function because the softmax over the values is insensitive to the global events larger constant. However,global and local model predictions differ when paired with polynomial or exponential complexity, as the global cost leads to a prediction of increasingly conservative intervening later in the trial (i.e. as the total event number increases the model becomes highly likely to do nothing), while the local cost depends dynamically on how much has happened recently.

Our primary class of models is based on the utility function specified in Equation 11. That is, one that is sensitive to both expected information and computational cost but differs in the measurement of this cost. To fit human decisions with these utilities we additionally added two components. We added a general bias against action $B(i^*) = [0, \dots, 1]$ capturing the fact that the vast majority of time windows do not contain an action.[10]

We additionally assumed stochasticity in participants' intervention choices captured by a softmax function (Luce, 1959) over the resultant values:

$$P(\text{choice} = i^*) = \frac{\exp\Big((\text{EIG}[i^*] - \omega\, \text{ECC}[i^*, c, f] + \theta B)/\tau\Big)}{\sum_{i' \in \mathbf{A}} \exp\Big((\text{EIG}[i'] - \omega\, \text{ECC}[i', c, f] + \theta B)/\tau\Big)} \tag{14}$$

The temperature parameter $\tau \in (0, +\infty)$ controls the spread of predictions across the response options. As $\tau \to +\infty$ model predictions become increasingly deterministic, while if $\tau \to 0$ the predictions become increasingly uniform. To investigate whether participants are sensitive to both expected information gain and expected computational cost, we also examined a pure information-driven model that removes ECC from Equation 14, and a strong baseline model that removes both ECC and EIG from Equation 14.

We use hold-one-device-out cross-validated log-likelihood as our primary measure of model fit but also include BIC for completeness and comparison with past work (Tauber et al., 2017). Our cross-validation scheme is conservative, since it requires a unified explanation for human data, in spite of different causal devices exhibiting very different characteristic dynamics. The temperature parameter τ and scaling parameters ω and θ were jointly fit using the `optim` function in R with BFGS method. We assumed a generic exponent of 2 in the polynomial complexity and a base of 2 for exponential complexity ($c = 2$) but we examine this assumption via a coarse grid search in Table A2.

Model fitting results

Results are shown in Table 4. Across both experiments, models that considered both expected information gain and inference cost outperformed pure information driven models. The best variant for both experiments was one that anticipated costs on the basis of a polynomial function of the expected local events. Models including both information and costs also better fit more individuals in both experiments than the other models we considered (87% people are fit by one of the cost-dependent models, 56% people are fit by the local-polynomial cost model specifically). Figure 15 gives an example from

[10] This is necessary to accommodate the fact that greedy-EIG underestimates the value of waiting when opportunities to intervene are finite. Conceptually, this is because it does not incorporate the expected utility of saving an action for use later.

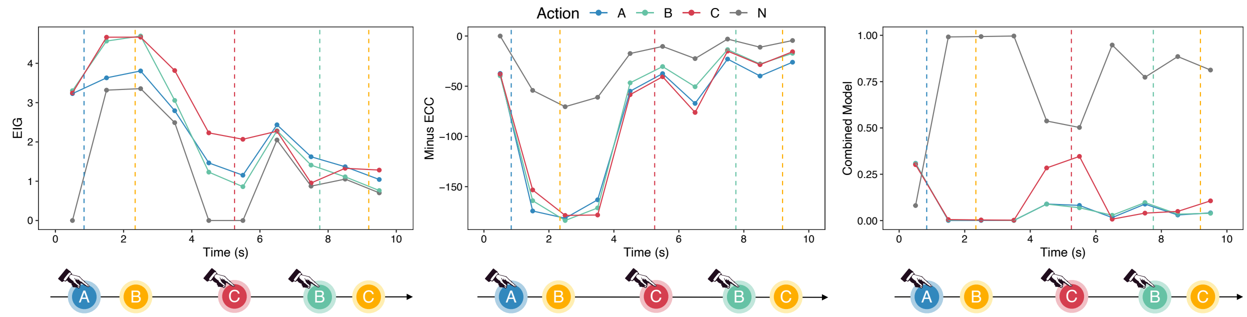

Figure 15

Example of real-time model prediction in Experiment 1. Events were generated by one of the participants in the reliable condition with a hidden ground truth of $A \rightarrow B \rightarrow C$ structure. Models take previous interventions and observations into consideration. Expected IG (information gain) and local polynomial cost are discounted by time. Parameters of the combined model (which also considers choices' baseline probability) are based on IG + local polynomial cost model of the individual. For interpretation of colors in this figure, the reader is referred to the online version of this article.

Experiment 1 in which the combination of expected information and cost in combination give a better account of participants' intervention choices than either does alone.

Model fitting discussion

The win for our bounded rational models over pure information seeking models supports our central claim that participants' interventions were shaped both by how much information they expected to gain and by how hard they expected to process potential future information. Note that while we do not present them for space reasons here, model variants sensitive to only cost but not information perform worse than all the models we present irrespective of how cost is calculated, always favoring waiting or blocking over activating components. However, a few more individuals in Experiment 2 were fit by the cost-free model according to BIC. This suggests that cost-free and cost-dependent models did not differ as much as in Experiment 1 when explaining human interventions. This might due to that the computational cost component of the model predicted learners should block fairly frequently, while participants blocked less often than predicted in general. We suspect that this is partly due to a preference for simplicity but in terms of strategy choice rather than evidence, with blocking strategies being intuitively more involved. Furthermore, our models so far only consider information gain and computational cost of the current intervention, while as discussed, people are likely to plan ahead when using blocking, for instance combining a block with a subsequent activation, which goes beyond the capability of our greedy models.

General Discussion

In a dynamically unfolding world, uncovering causal relationships requires processing continuous sensory information, tracking what happens before, during, and after ones actions. Compounding this, one must choose *where* to act, *how* to act, and *when* to do so. In this paper, we investigated human learning in a setting where learners intervened in a system to reveal what underlying causal structure produced the observed events unfolding in real time. We investigated what factors affected the inference process, and what strategies people used to choose and time their interventions. To answer these questions, we first established normative benchmarks for inference and for efficient intervention choice in our task, and then compared these against human behavior. We hypothesized that computational limitations, and a rational anticipation thereof, would play a role in shaping real time active learning. Thus, we endeavored to quantify the actual and anticipated computational cost of the evidence stream in our task to test whether this could help account for human judgments and intervention patterns.

Our main findings fall into two classes: (1) Insights about environmental factors that determine the success of active causal learning in continuous time and (2) insights about how people choose interventions to support their learning in continuous time.

Active causal learning performance. In terms of the former category, we found that people had more success identifying the structure of acyclic systems. The reliability of the delays between causes and effects had less impact on success that we might have expected based on past work (Bramley, Gerstenberg, Mayrhofer et al., 2018; Greville & Buehner, 2010), with unreliable delays having a sevenfold wider standard deviation yet appearing to moderately hamper inference only in cyclic devices and those involving 4 rather than 3 components. It could be that more extreme irregularity would have had a stronger effect on learning. On the other hand, recent work on observational causal learning has generally found *order* to be a stronger driver of structure judgments than precise delays (Bramley, Gerstenberg, Mayrhofer et al., 2018; Valentin et al., 2020). In this interactive setting, learners were able to test and retest the same component keeping an eye out for order reversals in its consequences. Unlike participants, our ideal observer was more accurate at identifying the structure of cyclic devices and more accurate again, when delays were also reliable. This serves to highlight the divergence between ideal and bounded learning. Our ideal observer could take advantage of large numbers of events to learn better and was sensitive to precise delay evidence, while humans are constrained by their information processing capacity. Participants' accuracy on different devices was inversely associated with the number of events occurring during the learning process, particularly with the peak event density during a trial. Both measures are conceptually related to computational demand of causal inference, since one should, ideally, consider all causal pathways that could link the events. In Experiment 2, where participants have more ability to control complexity using blocks, we still found the same relationship between event count and accuracy but found that other manipulated factors (delay reliability, structure cyclicity, structure nodes) no longer significantly affected accuracy.

Intervention decisions. In terms of the latter class — how people use interventions to support their learning in continuous time — we found that participants waited longer when the true structures contained cycles and also were locally reactive to the number of expected future events. Participants did not use blocks nearly as often as activations, but most participants did use them and did so particularly for excitable structures and at moments of high anticipated complexity. We distinguished tentatively between two uses of blocks: Resetting the system and performing Controlled tests that combine blocks with subsequent activations. We found that Resets, in particular, were used more often when the expected number of future events was high. Finally, our modeling confirmed that participants intervention decisions across both experiments were better fit by models sensitive to both information gain and computational cost.

Resource-rational active structure learning

The bounded nature of cognitive computation has long been discussed in relation to models of human learning (Anderson, 1990; Simon, 1982). While early research conceptualized the role of cognitive resources qualitatively, more recent studies have aimed to quantify cognitive costs and estimate boundedly rational norms (Griffiths et al., 2015; Lieder & Griffiths, 2020; Vul et al., 2014). Utility functions that combine both expected rewards and computational costs have been shown to better capture a variety of human behaviors including estimation (anchoring-and-adjustment, Dasgupta et al., 2017; Lieder et al., 2018), planning (Correa et al., 2020), information sampling (Petitet et al., 2021), decision making (Gershman, 2020), and communication (Hawkins et al., 2021). The current paper extends this line of research to the problem of real-time active causal learning. We showed that in addition to the standardly-considered exogenous costs of interventions (Coenen et al., 2015; Coenen, Ruggeri et al., 2019), people also care about the internal costs that arise from integrating different forms of causal interaction data. That is, learners were sensitive to the fact that information following an intervention has to be process-able to be useful.

Specifically, we found that, out of the measures we tried, a polynomial function of expected future events best captured the influence of complexity on intervention choice. While it would be premature to take this functional form as final, or to make a judgment about whether participants under- or over-anticipated the actual effect of complexity on their inference, we feel this reflects an intuitively sensible and plausible sensitivity to local events and the way in which inferential complexity compounds as the number of actual causal relata increase (Bramley, Dayan et al., 2017; Fernbach & Sloman, 2009; Van Rooij et al., 2019).

The issue of just how complexity scales raises a question as to what inference process learners actually used in this task. Although many papers, including this one, have laid out computational-level mechanisms of causal structure induction (Griffiths &

Tenenbaum, 2009; Pearl, 2000; Rottman & Hastie, 2014), they are typically intractable, requiring a run time that scales often far worse than exponentially (c^n) in the number of relata. As has been argued forcefully elsewhere (Van Rooij et al., 2019), this makes most computational-level models "non-starters" as process accounts of human inference in natural settings, since any plausible account will have to deal with more than a handful of components or events without requiring a time to compute that is beyond the lifespan of the organism (or even the universe). We note that the resource-rational framework adds another layer of computation, which is itself intractable. We use it here to establish that people are sensitive to information and computational cost but we do not provide a recipe for how learners anticipate these costs, given that this depends critically on their inferential processing. Human learning is necessarily more piecemeal and approximate and indeed, human responses are much noisier than our ideal observers'. There are some promising avenues for process accounts that can model aspects of this variability and noise. Simulation-based (Gerstenberg et al., 2021), summary statistics (Gong & Bramley, 2020), and incremental search (Bramley, Dayan et al., 2017) algorithms have all been proposed in recent years as aspects of how learners simplify and approximate solutions to structure learning.

When considering complexity, it is perhaps surprising that there was not more difference in performance between 3- and 4-variable problems since the latter involve an order of magnitude more hypotheses. However, this is in line with recent process level accounts. It has been argued that learners form one or a few hypotheses at a time (Bonawitz et al., 2014), or on subparts of the larger system (Davis et al., 2020; Fernbach & Sloman, 2009). These accounts are better able to scale up to inferences among more relata (Bramley, Dayan et al., 2017). Bramley, Dayan et al. (2017) show that in inference from interventions in covariation setting, people rely on sequential local changes to gradually update their beliefs to incorporate new evidence. Compared to maintaining a global prior, this localist approach may help people to deal with situations involving more than a handful of variables without invoking an exponential increase in computation or catastrophic loss of performance. Thus we conclude from this pattern, that however people manage the complexity of real time causal structure inference, they do so in such a way that they are affected by the number of events, but less by the number of components. Indeed, the run time of our our ideal observer was far more sensitive to the number of paths it had to evaluate per possible hypothesis than the number of hypotheses it evaluated.

Heuristics

One long-standing debate centers on the question of whether human intervention choice involves anticipation of information at all, or whether it relies on heuristics such as simple endorsement (Bramley et al., 2015), positive or confirmatory testing (Coenen et al., 2015; Steyvers et al., 2003). In the current setting, one reasonable heuristic might be to explore components until effects are discovered. If a component appears to produce multiple effects a learner can repeat-test it or probe the components at which the effects occur. In this way learners might follow a kind of positive testing strategy in which they focus their energies on components deemed to be root components so as to gather evidence "by making the machine go". This reflects the rationale behind positive testing strategy that has featured in the literature on atemporal active causal learning (Austerweil & Griffiths, 2011; Bramley, Dayan et al., 2017; Coenen et al., 2015; Klayman & Ha, 1989; Steyvers et al., 2003). However, unlike in the atemporal setting, there is lots that can be learned by repeatedly testing suspected root components in our experiments, meaning it is hard to distinguish whether the repeat selection of root components was driven by explicitly computing expected information gain or by following a simpler strategy such as combining random exploration with positive testing. Addtionally, it is possible that people choose when to intervene separately to where to intervene, for instance using current complexity as a way to decide when to perform one's next intervention and then selecting this without regard to anticipated complexity. To resolve these questions about psychological processing, future studies could set up continuous-time active causal learning scenarios that pit the predictions of heuristics against those of rational norms.

Continuous-time causal learning

While participants performed a handful of interventions per trial and experienced a handful of effects, this still constitutes a small amount of evidence compared to many contingency datasets used in causal inference contexts (Buehner et al., 2003; Griffiths & Tenenbaum, 2005). On the other hand, temporal order and delay enrich the information provided by each event allowing our idealized learning algorithms to draw strong and accurate conclusions from the data in most trials. We argue that thinking through the temporal micro-dynamics of causal mechanisms in small data contexts may better reflect everyday forms of causal reasoning than reasoning from large samples and statistically aggregated data, as everyday life does not generally provide these.

Some questions related to continuous-time causal learning remain for future research. One open question is where cyclic structures fit into the causal cognition picture. A full representation of a cyclic system seems to demand a temporal dimension and predictions are generally sensitive to the system's current state. Researchers have tried to model the change of continuous variables in continuous time (Davis et al., 2020; Soo & Rottman, 2018), while the current paper provides a new framework to model events in continuous time. In our experiments, participants performed well above chance for most connections in most cyclic structures without extra training, and could reliably determine if a structure contained a cycle even though they were less accurate in identifying the exact structure. This suggests that they can understand cyclic relationship relatively intuitively. There were two exceptions where people didn't do better than chance. These were (`Cyclic3` and `Cyclic7` in Figure 5 and 12). Both of these devices involved a component that is an output of a loop. Take `Cyclic3` ($A \leftrightarrow B \rightarrow C$) for example, the activation of B

tends to lead to activations of A and C very close in time, then another B and so on. In this case, we hypothesize that learners found it hard to tell whether the loop was perpetuated by B's or C's activation. An analogous ambiguity in reality could arise wherever it is unclear which events are pure effects (with no potential to control the system dynamics, such as symptoms of a disease) and which are constituent parts of the system's feedback loop (such as the pathogen). If one is interested in controlling a cyclic system, it is important to identify and act on components that are inputs to, or constituent parts of the feedback loop, rather than pure effects, since only by doing so can one nudge the system toward whatever state one wants it to take. Future studies might further use and develop our framework to have a closer look of cognitive processes in controlling systems involving cyclic causal relationships.

Another open question concerns the relationship between temporal and covariation-based causal learning. One possibility is that these depend on separate learning processes, but it also seems likely that there are points of overlap. For example, people may extract covariation information from continuous-time evidence through some process of abduction and distretization. Furthermore, interventions might help to create data that is more amenable to these forms of summarisation. Better understanding of human causal induction requires us to go beyond covariation-based causal Bayesian networks (Pearl, 2000), but this should not involve discarding their strengths and the insights they have provided in the search for a unified account for causal learning. Our current paradigm simplifies causes and effects to point events with no measurable duration. However, actual events are often extended in time in complex ways and many require reset or refractory period between occurrences. Therefore, it might be informative to also consider a setting in which causes must be reset or take time to recover to make this paradigm more comparable to the statistical-based causal learning.

Conclusion

Everyday experience is replete with events that reoccur, and where some of which are causally related in ways that allow us to predict, control, and make sense of what happens. While previous research on active causal learning has sidestepped the temporal dimension, in this paper we show that human learners are sensitive to time, not just in terms of how it impinges on what can be learned from evidence in principle, but also in terms of how it shapes the practicalities of interpreting that evidence as it arrives. Our experiments and modeling show that participants adapted their actions to the ongoing event dynamics during learning so as to strike a balance between expected information gain and anticipated inferential cost. These results contribute to our understanding of causal inference in continuous time, incorporate a new dimension to the study of human active learning and offer new directions for causal learning research.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.

Austerweil, J. L. & Griffiths, T. L. (2011). Seeking confirmation is rational for deterministic hypotheses. *Cognitive Science*, *35*(3), 499–526.

Bonawitz, E., 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.

Bramley, N. R., Dayan, P., Griffiths, T. L. & Lagnado, D. A. (2017). Formalizing neurath's ship: Approximate algorithms for online causal learning. *Psychological Review*, *124*(3), 301.

Bramley, N. R., Gerstenberg, T., Mayrhofer, R. & Lagnado, D. A. (2018). Time in causal structure learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(12), 1880–1910.

Bramley, N. R., Gerstenberg, T., Tenenbaum, J. B. & Gureckis, T. M. (2018). Intuitive experimentation in the physical world. *Cognitive Psychology*, *195*, 9–38.

Bramley, N. R., Lagnado, D. A. & Speekenbrink, M. (2015). Conservative forgetful scholars: How people learn causal structure through sequences of interventions. *Journal of Experimental Psychology: Learning, Memory, and 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 G. Gunzelmann, A. Howes, T. Tenbrink & E. J. Davelaar (Eds.), *Proceedings of the 39th annual conference of the cognitive science society* (pp. 150–155).

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, and Cognition*, *29*(6), 1119.

Buehner, M. J. & May, J. (2002). Knowledge mediates the timeframe of covariation assessment in human causal induction. *Thinking & Reasoning*, *8*(4), 269–295.

Burns, P. & McCormack, T. (2009). Temporal information and children's and adults' causal inferences. *Thinking & Reasoning*, *15*(2), 167–196.

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.

Christiansen, M. H. & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, 1–72.

Coenen, A., Nelson, J. D. & Gureckis, T. M. (2019). Asking the right questions about the psychology of human inquiry: Nine open challenges. *Psychonomic Bulletin & Review*, *26*(5), 1548–1587.

Coenen, A., Rehder, B. & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology*, *79*, 102–133.

Coenen, A., Ruggeri, A., Bramley, N. R. & Gureckis, T. M. (2019). Testing one or multiple: How beliefs about sparsity affect causal experimentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(11), 1923–1941.

Correa, C. G., Ho, M. K., Callaway, F. & Griffiths, T. L. (2020). Resource-rational task decomposition to minimize planning costs. In S. Denison, M. Mack, Y. Xu & B. Armstrong (Eds.), *Proceedings of the 42th annual conference of the cognitive science society* (pp. 2974–2980).

Dasgupta, I., Schulz, E. & Gershman, S. J. (2017). Where do hypotheses come from? *Cognitive Psychology*, *96*, 1–25.

Davis, Z., Bramley, N. R. & Rehder, B. (2020). Causal structure learning in continuous systems. *Frontiers in Psychology*, *11*, 244.

Derringer, C. & Rottman, B. M. (2018). How people learn about causal influence when there are many possible causes: A model based on informative transitions. *Cognitive Psychology*, *102*, 41–71.

Fernbach, P. M. & Sloman, S. A. (2009). Causal learning with local computations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 678–693.

Frank, S. L. (2013). Uncertainty reduction as a measure of cognitive load in sentence comprehension. *Topics in Cognitive Science*, *5*(3), 475–494.

Galton, F. (1890). Mental tests and measurements. *Mind*, *15*(59), 373–381.

Gershman, S. J. (2020). Origin of perseveration in the trade-off between reward and complexity. *Cognition*, *204*, 104394.

Gerstenberg, T., Bechlivanidis, C. & Lagnado, D. A. (2013). Back on track: Backtracking in counterfactual reasoning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *35*(35).

Gerstenberg, T., Goodman, N. D., Lagnado, D. A. & Tenenbaum, J. B. (2021). A counterfactual simulation model of causal judgments for physical events. *Psychological Review*, *128*(5), 936–975.

Gilbert, E. (1966). Information theory after 18 years: After years of theoretical development, information theory may fulfill its engineering promise. *Science*, *152*(3720), 320–326.

Gong, T. & Bramley, N. R. (2020). What you didn't see: Prevention and generation in continuous time causal induction. In S. Denison, M. Mack, Y. Xu & B. Armstrong (Eds.), *Proceedings of the 42th annual conference of the cognitive science society* (pp. 2908–2914).

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), 620–629.

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.

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*(2), 217–229.

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*(4), 661–716.

Guillory, A. (2012). *Active learning and submodular functions*. University of Washington.

Hagmayer, Y. & Waldmann, M. R. (2002). How temporal assumptions influence causal judgments. *Memory & Cognition*, *30*(7), 1128–1137.

Halpern, J. Y. (2016). *Actual causation*. MIT Press.

Hawkins, R. D., Gweon, H. & Goodman, N. D. (2021). The division of labor in communication: Speakers help listeners account for asymmetries in visual perspective. *Cognitive Science*, *45*(3), e12926.

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), 451–476.

Klayman, J. & Ha, Y.-w. (1989). Hypothesis testing in rule discovery: Strategy, structure, and content. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 596–604.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, *96*(4), 674–689.

Kuhn, D. & Brannock, J. (1977). Development of the isolation of variables scheme in experimental and" natural experiment" contexts. *Developmental Psychology*, *13*(1), 9–14.

Lagnado, D. A. & Sloman, S. (2002). Learning causal structure. In W. Gray & C. Schunn (Eds.), *Proceedings of the 24th annual meeting of the cognitive science society* (pp. 560–565).

Lagnado, D. A. & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 451–460.

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y. & Sloman, S. A. (2007). Beyond covariation. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 154–172). Oxford University Press.

Lai, L. & Gershman, S. J. (2021). Policy compression: An information bottleneck in action selection. *The Psychology of Learning and Motivation*, 195–227.

Lieder, F. & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, 1–60.

Lieder, F., Griffiths, T. L., Huys, Q. J. & Goodman, N. D. (2018). The anchoring bias reflects rational use of cognitive resources. *Psychonomic Bulletin & Review*, *25*(1), 322–349.

Lowder, M. W., Choi, W., Ferreira, F. & Henderson, J. M. (2018). Lexical predictability during natural reading: Effects of surprisal and entropy reduction. *Cognitive Science*, *42*, 1166–1183.

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.

Luce, R. D. (1959). *Individual choice behavior*. Wiley.

Malthus, T. R. (1872). *An essay on the principle of population*. Reeves & Turner.

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

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.

Meng, Y., Bramley, N. & Xu, F. (2018). Children's causal interventions combine discrimination and confirmation. In T. Rogers, M. Rau, X. Zhu & C. W. Kalish (Eds.), *Proceedings of the 40th annual conference of the cognitive science society*.

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

Nelson, J. D. (2005). Finding useful questions: On bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, *112*(4), 979–1000.

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.

Pacer, M. & Griffiths, T. L. (2012). Elements of a rational framework for continuous-time causal induction. In N. Miyake, D. Peebles & R. P. Cooper (Eds.), *Proceedings of the 34th annual conference of the cognitive science society* (pp. 833–838).

Pacer, M. & Griffiths, T. L. (2015). Upsetting the contingency table: Causal induction over sequences of point events. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings & P. P. Maglio (Eds.), *Proceedings of the 37th annual conference of the cognitive science society* (pp. 1805–1810).

Pearl, J. (2000). *Causality*. Cambridge University Press (2009 reprint).

Petitet, P., Attaallah, B., Manohar, S. G. & Husain, M. (2021). The computational cost of active information sampling before decision-making under uncertainty. *Nature Human Behaviour*, 1–12.

Pierce, J. R. & Karlin, J. E. (1957). Reading rates and the information rate of a human channel. *Bell System Technical Journal*, *36*(2), 497–516.

Rehder, B. (2014). Independence and dependence in human causal reasoning. *Cognitive Psychology*, *72*, 54–107.

Rehder, B. (2017). Reasoning with causal cycles. *Cognitive Science*, *41*, 944–1002.

Rottman, B. M. & Hastie, R. (2016). 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*(1-2), 93–125.

Rottman, B. M. & Hastie, R. (2014). Reasoning about causal relationships: Inferences on causal networks. *Psychological Bulletin*, *140*(1), 109–139.

Schulz, L. E., Gopnik, A. & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science*, *10*(3), 322–332.

Settles, B. (2009). *Active learning literature survey*. Technical Report University of Wisconsin-Madison.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D. & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*, 99–124.

Simon, H. A. (1982). *Models of bounded rationality: Empirically grounded economic reason*. MIT press.

Sloman, S. A., Love, B. C. & Ahn, W.-K. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, *22*(2), 189–228.

Sobel, D. M. & Kushnir, T. (2006). The importance of decision making in causal learning from interventions. *Memory & Cognition*, *34*(2), 411–419.

Soo, K. W. & Rottman, B. M. (2018). Causal strength induction from time series data. *Journal of Experimental Psychology: General*, *147*(4), 485–513.

Spirtes, P., Glymour, C. N., Scheines, R. & Heckerman, D. (2000). *Causation, prediction, and search* (2nd ed.). MIT press.

Stephan, S., Mayrhofer, R. & Waldmann, M. R. (2020). Time and singular causation–a computational model. *Cognitive Science*, *44*(7), e12871.

Stephan, S., Tentori, K., Pighin, S. & Waldmann, M. R. (2021). Interpolating causal mechanisms: The paradox of knowing more. *Journal of Experimental Psychology: General*, *150*(8), 1500–1527.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E.-J. & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*(3), 453–489.

Tauber, S., Navarro, D. J., Perfors, A. & Steyvers, M. (2017). Bayesian models of cognition revisited: Setting optimality aside and letting data drive psychological theory. *Psychological Review*, *124*(4), 410.

Tong, S. & Koller, D. (2001). Active learning for structure in bayesian networks. *International Joint Conference on Artificial Intelligence*, *17*(1), 863–869.

Valentin, S., Bramley, N. R. & Lucas, C. G. (2020). Learning hidden causal structure from temporal data. In S. Denison, M. Mack, Y. Xu & B. Armstrong (Eds.), *Proceedings of the 42th annual conference of the cognitive science society* (pp. 1906–1912).

Van Rooij, I., Blokpoel, M., Kwisthout, J. & Wareham, T. (2019). *Cognition and intractability: A guide to classical and parameterized complexity analysis*. Cambridge University Press.

Verghese, P. & Pelli, D. G. (1992). The information capacity of visual attention. *Vision Research*, *32*(5), 983–995.

Vul, E., Goodman, N., Griffiths, T. L. & Tenenbaum, J. B. (2014). One and done? optimal decisions from very few samples. *Cognitive Science*, *38*(4), 599–637.

White, P. A. (2006). How well is causal structure inferred from cooccurrence information? *European Journal of Cognitive Psychology*, *18*(03), 454–480.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, *27*(2), 172–223.

Appendix

Table A1

Predictions of human causal learning accuracy using different window sizes of local events.

Window	β	t	Sig.	95% CI
Experiment 1				
1 s	-0.28	7.77	$< .001$	[-0.35, -0.20]
2 s	-0.23	6.38	$< .001$	[-0.31, -0.16]
3 s	-0.24	6.58	$< .001$	[-0.32, -0.17]
4 s	-0.24	6.29	$< .001$	[-0.32, -0.16]
5 s	-0.24	6.27	$< .001$	[-0.32, -0.16]
6 s	-0.23	5.97	$< .001$	[-0.32, -0.15]
7 s	-0.23	6.00	$< .001$	[-0.31, -0.15]
Experiment 2				
1 s	-0.09	3.55	$< .001$	[-0.14, -0.04]
2 s	-0.12	4.54	$< .001$	[-0.18, -0.07]
3 s	-0.14	5.12	$< .001$	[-0.20, -0.08]
4 s	-0.12	4.20	$< .001$	[-0.17, -0.06]
5 s	-0.12	4.27	$< .001$	[-0.18, -0.06]
6 s	-0.12	4.46	$< .001$	[-0.18, -0.07]
7 s	-0.11	4.02	$< .001$	[-0.17, -0.06]

Table A2

Intervention Model Fits of Exponents or Base Parameters in Polynomial- or Exponential-costs.

			Experiment 1			
	CV	BIC	$c_{BIC/CV}$	τ	ω	θ
Global cost						
EIG-ECC$_{\text{Polynomial}}$	-16451	32882	1.8	2.55	1.03×10^{-2}	8.52
EIG-ECC$_{\text{Exponential}}$	-16550	33067	1.8/2.4	2.79	1.44×10^{-10}	10.00
Local cost						
EIG-ECC$_{\text{Polynomial}}$	-16378	32743	1.4/1.6	1.80	6.22×10^{-2}	6.04
EIG-ECC$_{\text{Exponential}}$	-16556	33083	1.8	2.79	1.39×10^{-10}	10.02
			Experiment 2			
	CV	BIC	$c_{BIC/CV}$	τ	ω	θ
Global cost						
EIG-ECC$_{\text{Polynomial}}$	-40504	80901	3	1.96	4.40×10^{-6}	8.06
EIG-ECC$_{\text{Exponential}}$	-40503	80895	1.18/1.2	1.97	7.31×10^{-11}	8.09
Local cost						
EIG-ECC$_{\text{Polynomial}}$	-40462	80815	2	1.86	1.59×10^{-3}	7.63
EIG-ECC$_{\text{Exponential}}$	-40500	80896	3.2/1.18	1.97	8.65×10^{-11}	8.10

Note: The base and exponent parameter c was fit by grid search. We searched in (1,4), using a step of 0.002 for the range [1.002,1.018], a step of 0.02 for the range [1.02,1.18] and a step size of 0.2 thereafter. These steps approximate log-uniform intervals so are suitable for fitting a parameter bounded at the lower end but not at the higher end. Other parameters were fitted using BFGS given a fixed c. We reported two c with BIC and CV deviated in their results.