



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Active causal structure learning in continuous time

**Citation for published version:**

Gong, T, Gerstenberg, T, Mayrhofer, R & Bramley, NR 2023, 'Active causal structure learning in continuous time', *Cognitive Psychology*, vol. 140, 101542. <https://doi.org/10.1016/j.cogpsych.2022.101542>

**Digital Object Identifier (DOI):**

[10.1016/j.cogpsych.2022.101542](https://doi.org/10.1016/j.cogpsych.2022.101542)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Cognitive Psychology

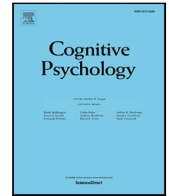
**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.





# Active causal structure learning in continuous time<sup>☆</sup>

Tianwei Gong<sup>a,\*</sup>, Tobias Gerstenberg<sup>b</sup>, Ralf Mayrhofer<sup>c</sup>, Neil R. Bramley<sup>a</sup>

<sup>a</sup> Department of Psychology, University of Edinburgh, United Kingdom

<sup>b</sup> Department of Psychology, Stanford University, United States of America

<sup>c</sup> Department of Psychology, University of Göttingen, Germany

## ARTICLE INFO

### Keywords:

Causal learning  
Intervention  
Time  
Resource rationality  
Causal cycles

## ABSTRACT

Research on causal cognition 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 they intervene 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, participants' intervention choices strike a balance between maximizing expected information and minimizing inferential complexity. 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.

## 1. 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 through *interventions* (Pearl, 2000) — actions that manipulate the environment in ways that reveal what causes what and distinguish spurious correlations from genuine causal relationships. However, learning causal structure in general, and selecting interventions in particular, are computationally challenging problems even under idealized conditions (Bramley, Dayan, Griffiths, & Lagnado, 2017). In everyday life, this challenge is compounded by the need to interact with the causal environment in real time, bringing computational constraints to the fore (Griffiths, Lieder, & Goodman, 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 about how the device works. 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; Simon, 1982) that serve as a guideline for our computational modeling framework. We then investigate human active learning about a range of acyclic and cyclic causal devices in two experiments. We analyze participants'

<sup>☆</sup> We thank Marc Buehner and two anonymous reviewers for many helpful comments. 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 et al., 2017). TiG is supported by a University of Edinburgh PPLS Scholarship, United Kingdom. RM is supported by a Deutsche Forschungsgemeinschaft Research Grant, Germany (WA 621/24-1). NB is part supported by a EPSRC New Investigator Grant, United Kingdom (EP/T033967/1).

\* Correspondence to: Department of Psychology, University of Edinburgh, 7 George Square, EH8 9JZ, Edinburgh, Scotland, United Kingdom.

E-mail address: [tia.gong@ed.ac.uk](mailto:tia.gong@ed.ac.uk) (T. Gong).

<https://doi.org/10.1016/j.cogpsych.2022.101542>

Received 14 January 2022; Received in revised form 7 December 2022; Accepted 11 December 2022

Available online 29 December 2022

0010-0285/© 2022 The Author(s).

Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Published by Elsevier Inc. This is an open access article under the CC BY license

causal judgments and intervention patterns both descriptively and through comparison with a range of models. We contrast an unbounded computational account that optimizes expected information density of its interactions with the devices with bounded models that balance information and inferential complexity. Finally, we discuss the broader implications of this perspective on accounts of human learning.

### 1.1. 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 food delivery 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 beliefs about causal mechanisms and affordances (Bramley, Dayan, et al., 2017; Griffiths & Tenenbaum, 2009; Lagnado & Sloman, 2002; Lucas & Griffiths, 2010; Meder, Gerstenberg, Hagmayer, & Waldmann, 2010; Rehder, 2014; Rottman & Hastie, 2016; Schulz, Gopnik, & Glymour, 2007; Sobel & Kushnir, 2006; Stephan, Tentori, Pighin, & Waldmann, 2021; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003). While the particulars of these studies are diverse, many share a core set of properties illustrated in Fig. 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 are binary with one value construed as a variable being “present” or “active”, and the other state as “absent”, or “inactive”.

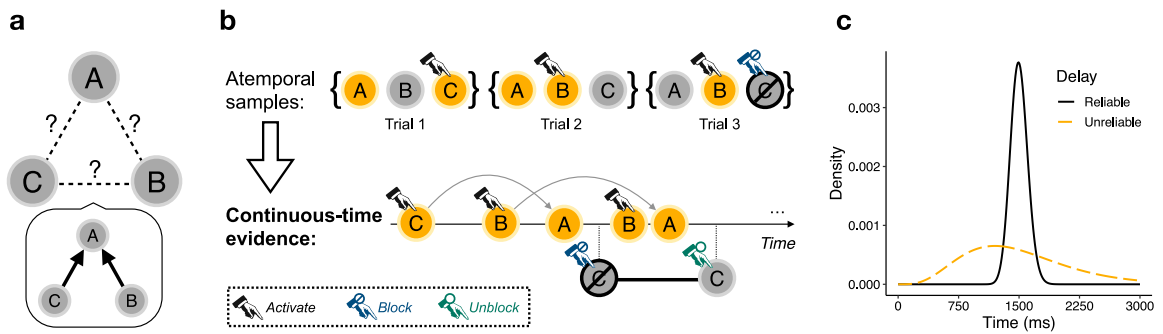
However, when only covariation data of a set of variables is available, observational samples are insufficient to uniquely reveal structure (Pearl, 2000; Spirtes, Glymour, Scheines, & Heckerman, 2000). For example, if a learner observes two variables co-occurring, such that when one is active (or inactive) the other one tends to be active (or inactive) too, they cannot tell if one is causing the other or if they share an 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 affects the rest of the system. A number of experiments have allowed participants to perform such interventions in order to support their learning (Bramley, Dayan, et al., 2017; Bramley, Lagnado, & Speekenbrink, 2015; Coenen, Rehder, & Gureckis, 2015; Lagnado & Sloman, 2002; Steyvers et al., 2003).

Studies have shown that well chosen interventions can speed up learning, allowing learners to target their uncertainty and quickly narrow in on the true model. However, poorly chosen interventions can be worse than random actions or passive observations (Settles, 2009). In the covariation-data setting, adults and children have been found to be 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, Bramley, Frosch, Patrick, & Lagnado, 2016; Meng, Bramley, & Xu, 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 inferring the true causal structure that gave rise to the data (Nelson, 2005). A number of experiments have demonstrated broad alignment with these norms in both adults and children, but also departures from the normative predictions which suggest that process-level considerations are important for fully characterizing people’s inferences (Bramley, Dayan, et al., 2017; Bramley et al., 2015). For example, people often chose an intervention that is expected to confirm or refute a currently-favored hypothesis rather than one that provides more information about the full hypothesis space (Coenen et al., 2015; Klayman & Ha, 1989; Meng et al., 2018; Steyvers et al., 2003). People sometimes also rely on 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 this strategy turns out to be suboptimal in certain environments (Bramley, Jones, Gureckis, & Ruggeri, 2022; Coenen, Ruggeri, Bramley, & Gureckis, 2019). Finally, adults choose interventions adaptively, taking into account environmental factors, such as time pressure, as well as whether a strategy was informative in the past (Coenen et al., 2015).

### 1.2. What prior work has neglected

Previous work on causal learning has largely focused on situations that mimic idealized laboratory conditions. In these studies, participants perform interventions in a discrete trial-by-trial manner, and the values of all variables are revealed all at once. In this way, participants are invited to generate and reason from a series of independent observations. Fig. 1b illustrates an example of this *atemporal* evidence, generated from interventions and subsequent observations of the variables in a stochastic system (see Bramley et al., 2015; Coenen et al., 2015, for example). Information arrives in three independent trials in the form of variable states (yellow = present or active; gray = absent or inactive) conditional on interventions (i.e. variables fixed on or blocked off by the learner).

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 Fig. 1b we see that, conditional on an intervention that activates *C*, *A* activated and *B* did not. This can be written as  $\{A = 1, B = 0\} \text{Do}[C = 1]$ , where 1 indicates a variable was active, 0 indicates it was inactive and  $\text{Do}[\dots]$  indicates a variable was fixed through intervention and thereby disconnected from its normal causes on this trial. Interventions can target multiple variables. For example, in Trial 3, both *B* and *C* are manipulated as *B* is activated and *C* is blocked (fixed to be inactive). In this kind of task, ideal inference and intervention selection are well-understood computationally, facilitating



**Fig. 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 possible corresponding continuous-time samples. Yellow indicates a component activated. In the continuous-time setting, interventions activate components in real time and effects may occur intermingled on the timeline. Arrows indicate the underlying generative process unbeknownst to the learner. (c) Gamma density distributions under reliable vs. unreliable causal delays in the current experiments. Both probability distributions have a mean of 1.5 s with different standard deviations (0.1 s for reliable and 0.7 s for unreliable delays). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

comparison between behavior and rational norms (e.g. Rottman & Hastie, 2014). However, this task setup differs in several respects from the causal learning and reasoning problems people face in daily life when they (1) take into account temporal information, (2) deal with evidence that is interdependent, and (3) encounter causal learning problems when the underlying causal mechanism may be cyclic.

### 1.2.1. Time

Most previous studies removed temporal information, including the order of events and the delays between them. For example, Coenen et al. (2015) described a cover story of computer-chip systems where the causal relationships are the passage of electrical current from the energy source to 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 inferring causal relationships. The notion that causes must precede their effects is foundational to the concept of causation (Burns & McCormack, 2009; Lagnado, Waldmann, Hagmayer, & Sloman, 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 would require an effect to have occurred at the same time as its cause (Bramley, Gerstenberg, Mayrhofer, & Lagnado, 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 as when they conform to prior or mechanistic expectations (Buehner & May, 2004; Buehner & McGregor, 2006; Hagmayer & Waldmann, 2002). For example, Buehner and McGregor (2006) found that participants gave higher causal judgments about the insertion of a ball turning on a light on a physical apparatus when the light came on after a few seconds rather than instantly, if they were aware it took time for the ball to roll through the apparatus and reach the light switch. A separate line of work has studied inference and representation of continuous variables in continuous time (Davis, Bramley, & Rehder, 2020; Soo & Rottman, 2018). However, temporal information is yet to be examined in the context of active causal learning.

### 1.2.2. 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, Cheng, & Clifford, 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, Sobel, Schulz, & Glymour, 2001; Lucas, Bridgers, Griffiths, & Gopnik, 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 wondering why your puppy is unusually excited one evening. You consider two candidate hypotheses: Perhaps his elevated mood is due to a new variety of dog food you fed him at 5pm, or 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 having smelt the flowers. A poor approach to resolving the question would be to always feed him beside the flower bed. It would be better 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 everyday life 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, many participants judged causal relationships by how values changed relative to their state on the preceding observation, rather than treating the samples as independent (see also [Derringer & Rottman, 2018](#)). This suggests that when evidence arrives over time, people strongly assume 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.

### 1.2.3. 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 the formalism of acyclic causal Bayesian networks (see [Rottman & Hastie, 2014](#), for a review). The continuous-time setting enables us to investigate cyclic causal relationships. A causal mechanism is cyclic if it has at least one component whose descendants include itself ([Pearl, 2000](#)). This means that the components that form part of the cycle, or outputs from it, may occur in repeated alternating fashion (e.g. a bidirectional connection  $A \leftrightarrow B$  could generate a sequence of events  $A, B, A, B, A, \dots$ ). Many causal processes in the natural world are cyclic ([Malthus, 1872](#)), and people frequently report causal beliefs that include cyclic relationships when allowed to do so in experiments ([Kim & Ahn, 2002](#); [Nikolic & Lagnado, 2015](#); [Rehder, 2017](#); [Sloman, Love, & Ahn, 1998](#)), making this an important aspect of causal cognition to study.

### 1.3. The current paradigm

*The learning problem.* Departing from the atemporal setting, we focus on what people can learn from interventions and observations of events within a single continuous timeline. We study a setting in which effect events follow their causes with some stochastic but predictable delay.

This causally-connected-point-event setting has been used in a number of recent studies of temporal causal reasoning. It rests on a firm mathematical foundation that supports normative inferences from temporal information to causal structure. [Greville and Buehner \(2010\)](#), [Griffiths and Tenenbaum \(2009\)](#) first demonstrated that people can infer how pairs of variables affect one another from observing sequences of point events. [Pacer and Griffiths \(2012, 2015\)](#) developed a model that infers causal relationships based on the occurrence of putative cause events that influence the rate at which the relevant effect events occur over time. [Bramley, Gerstenberg, Mayrhofer, and Lagnado \(2018\)](#) built models that combined hard order constraints with soft delay expectations to best capture structure judgments: Even when order information was fixed, participants were still sensitive to the variation in inter-activation delays between events and used it to distinguish between certain causal structures.

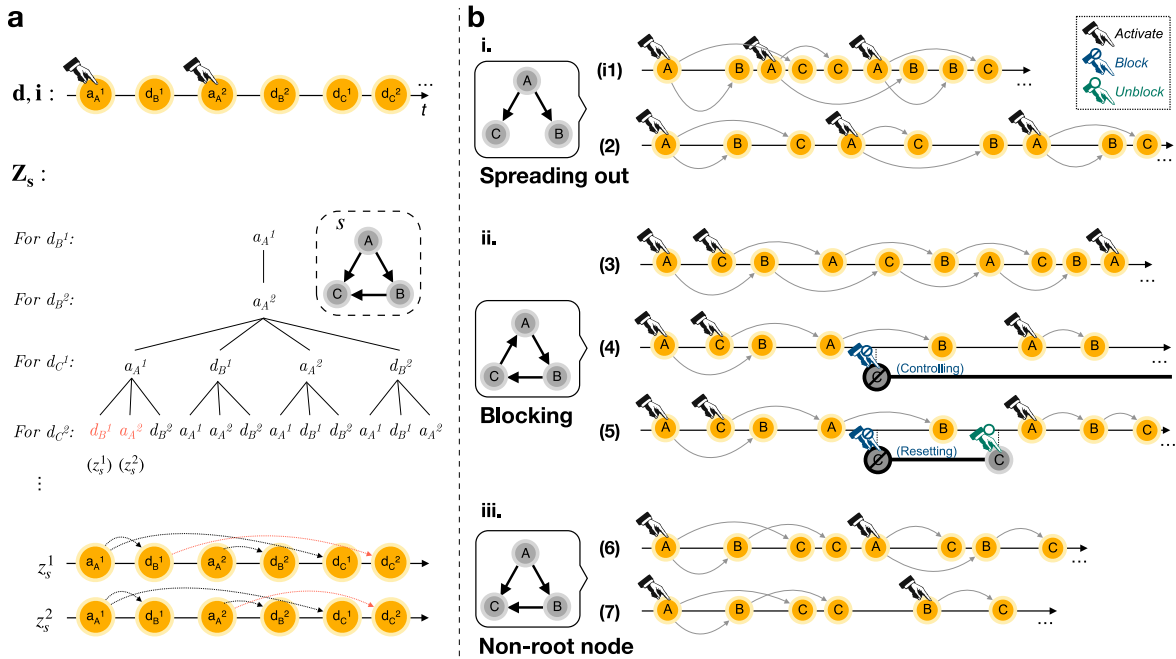
More recently, research has focused on so-called “actual causation” ([Halpern, 2016](#)): The question of which out of multiple candidate events actually brought about the outcome ([Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2021](#); [Stephan, Mayrhofer, & Waldmann, 2020](#)). Using key ideas from the “actual causation” literature, recent work has looked at how causal structures can be identified from temporal information by considering the different possible causal pathways that could have produced the observed events conditional on different underlying causal mechanisms ([Bramley, Gerstenberg, Mayrhofer, & Lagnado, 2018](#); [Gong & Bramley, 2020](#); [Valentin, Bramley, & Lucas, 2020](#)). We follow this approach, using Gamma distributions to model the distribution of causal delays exhibited by a particular causal component of a device across instances (see [Fig. 1b](#)). Gamma distributions define a probability density over  $(0, +\infty)$  via a shape parameter  $\alpha$ , and rate parameter  $\beta$  allowing for a variety of causal delay distributions with differing means and more or less variability (see [Fig. 1c](#)).

While temporal information was key to how the evidence was presented in the studies above, the data was not fully continuous in the sense in participants experiences were still broken into separate independent episodes. For example, in order to set things in motion in [Bramley, Gerstenberg, Mayrhofer, and Lagnado \(2018\)](#), 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. A fully continuous setting makes it more difficult to figure out what caused what because any given event might be attributed to an earlier-occurring event and might have its own effects that are still to occur.

*Activating and blocking interventions in time.* In our experiments we will allow learners to intervene on the causal system in two ways:

1. By *activating* components, thus potentially setting in motion a new sequence of events.
2. By *blocking* components, thus preventing that component both from 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  $\text{Do}[X = 1]$  or off  $\text{Do}[X = 0]$  in the atemporal 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. Thus, activation is better thought of as a shock to the system than as a form of graph surgery.



**Fig. 2.** Sketch of ideal observer inference algorithm and approaches to minimizing complexity. (a) Ideal Bayesian inference considers each possible structure hypothesis  $s$  and every possible causal path  $z_s^i$  that could describe how that structure produced the observations. The number of possible paths grows rapidly in the number of “nearby” events as illustrated with an example recursion tree showing all twelve paths connecting events  $a_A^1 \dots d_C^2$  conditional on the structure  $C \leftarrow A \rightarrow B \rightarrow C$  ( $d_{\text{component}}^{(\text{index})}$ : activating interventions;  $a_{\text{component}}^{(\text{index})}$ : effect events, see Appendix A for a full description of notation). Two paths  $z_s^1$  and  $z_s^2$  were further displayed in a timeline format with arrows showing the hypothesized generative process and red arrows in particular highlighting the different delay implications. (b) Three examples of interventional strategies that help reduce the inferential cost of processing generated evidence. Sequences (2), (4), (5), (7) are less complex to process than Evidence (1), (3), (6).

On the other hand, *blocking* actions do exemplify the “graph surgery” property in Pearl’s sense. They disconnect the blocked component from its normal causes until it is unblocked again (Fig. 1b). In the atemporal setting, blocking is essential for discriminating between certain structures (Bramley et al., 2015; McCormack et al., 2016; Schulz et al., 2007). For example, turning on a single component (i.e.  $Do[A = 1]$ ,  $Do[B = 1]$  or  $Do[C = 1]$ ) generates a similar pattern of dependence under a  $A \rightarrow B \rightarrow C$  chain and a  $C \leftarrow A \rightarrow B \rightarrow C$  fully connected structure — activating A affects B and C, activating B affects C but not A, activating C neither affects A nor B. This makes it difficult and inefficient to distinguish these structures based on activating interventions alone and impossible in deterministic settings. 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 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 system 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 useful for reducing computational complexity and ambiguity of parsing the consequent event sequences. The learner can use blocking to reduce the event numbers, or remove a component from consideration, while still making remaining events informative (see the section below for examples of two relevant strategies).

**1.4. Cognitive resource limitations**

In the atemporal 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 prompts, bringing ones inferential tools to bear quickly and efficiently without the luxury of time to backtrack.

We lay out the how an ideal Bayesian observer learns the causal structure with temporal evidence in Appendix A. This shows that the amount of computation needed to process the evidence compounds rapidly as more events occur. An ideal learner needs to consider all the plausible pathways through which a particular causal structure might have produced an observed pattern of events (Halpern, 2016). For example, consider intervening once each on A and B and then observing two subsequent activations of C. To calculate the overall likelihood that a “collider” structure  $A \rightarrow C \leftarrow B$  could have produced this pattern, we would need to take into account two possibilities (1) that A produced the first activation of C and B the second one, or (2) that B caused the first one, and

A the second one. However, there will generally be far more than two such possibilities. Fig. 2a shows a more complex example in which twelve possible causal paths could link six events for a single causal structure. For a handful more events the number of paths can easily grow into the millions.

Challenging such a naive idealized account of causal inference is the basic fact that human minds are bounded in their capacity to compute and store information. Human reasoning and decision-making necessarily deviates from such intractable, computational-level ideals (Anderson, 1990; Simon, 1982). Given the computational complexity and conceptual centrality of structure learning in cognition, we expect computational costs to play a large role when structure inference must take place in real time (Christiansen & Chater, 2016). Several process-level proposals have been explored in the literature as candidates for how people approximate normative structure inference. People may only consider a few sampled hypotheses (Bonawitz, Denison, Gopnik, & Griffiths, 2014), incrementally adjust a focal hypothesis to accommodate new evidence (Bramley, Dayan, et al., 2017; Davis et al., 2020; Fernbach & Sloman, 2009), rely on recent evidence (Bramley, Dayan, et al., 2017; Bramley et al., 2015), rely on summary statistics (Gong & Bramley, 2020, 2022; Ullman, Stuhlmüller, Goodman, & Tenenbaum, 2018), or on simple heuristics such as equating temporal order with causal order (Bramley, Gerstenberg, Mayrhofer, & Lagnado, 2018; Bramley, Mayrhofer, Gerstenberg, & Lagnado, 2017; Burns & McCormack, 2009).

While we expect some combination of the above ideas to be in play in how participants solve our task, we here explore a complementary facet of causal learning, the active gathering of evidence through interventions to support the inference process. We ask whether people time and target their interventions so as to manage the inferential complexity of parsing resultant evidence, while still producing informative evidence overall. Fig. 2b illustrates this idea by displaying three potential intervention strategies for managing evidential complexity. In the first example, the ground truth is  $B \leftarrow A \rightarrow C$ . An unbounded ideal learner learns about as much from Evidence 1 as from Evidence 2. However, if we assume that the ability to process evidence is a function of its complexity and that this is related to the density of the events being reasoned about, then it is clear that Evidence 2 is the more useful for a bounded learner. Here, the events are better separated, and so there is much less ambiguity about the plausible causes of each token event, therefore less need to engage in costly averaging over many potential causal pathways under each structure hypothesis.

Bounded learners may also choose to *block* components of a system to make the event stream manageable for reasoning. As shown in Evidence 3 of Fig. 2b, having performed two activating interventions in a cyclic system, a learner may experience a confusing pattern of parallel excitation. Evidence 4 shows how this can be avoided using a “controlled” testing strategy that blocks a component before activating another. This approach allows a learner to isolate a subsystem of a larger system. This controlled test means fewer interpretations of the evidence need be considered. For instance, in Evidence 3 remaining events are straightforwardly indicative of the substructure linking the unblocked components  $A$  and  $B$ . Well-timed blocks might also be used to impose pseudo-independence and trial-like structure within a continuous interaction. For example, as shown in Evidence 5, one might block, wait, then unblock components to “reset” a system, preventing any ongoing activity from complicating the inference process going forward.

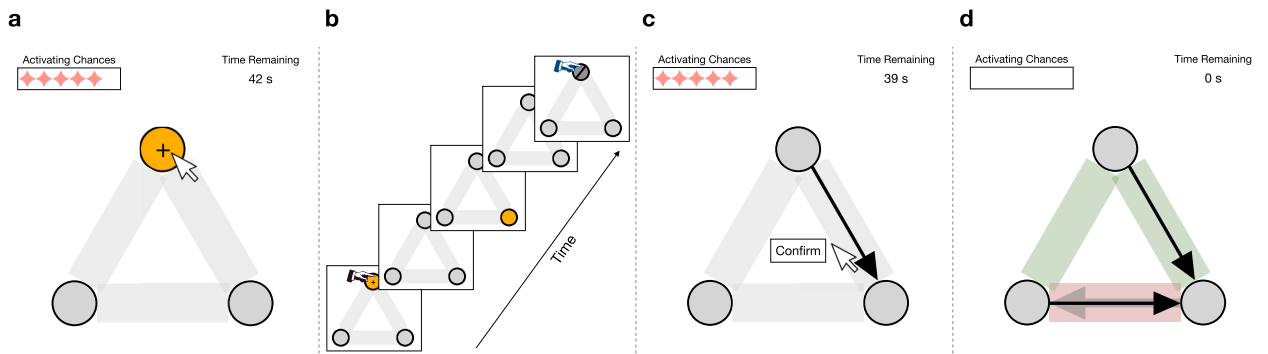
Finally, although activating a suspected root component (here  $A$ ) tends to produce more evidence about a causal structure than activating its suspected tail nodes (Evidence 6), bounded learners might sometimes avoid root components. For instance, if primarily interested in understanding a presumed downstream subpart of a complicated causal mechanism, one might intervene locally to avoid extraneous events and activity (Evidence 7). Note that those considerations are not independent. Spreading out interventions in time, for example, requires the learner to wait for the system to calm down. The same could be accomplished by blocking-and-unblocking a component to reset the system.

Finding ways to balance informativeness and complexity in generating evidence is conceptually related to the notion of bounded rationality (Anderson, 1990; Simon, 1982). The basic idea is that human minds have evolved or discovered solutions that trade off efficiently between the costs of computation and its rewards in greater accuracy or performance. In particular, one can incorporate computational costs into a solution space formally with a *resource rationality analysis* (see Griffiths et al., 2015; Lieder & Griffiths, 2020; Shenhav et al., 2017, for review). This has suggested that a number of decision making phenomena classically 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 (Callaway et al., 2022; Dasgupta, Schulz, & Gershman, 2017; Hawkins, Gweon, & Goodman, 2021; Lai & Gershman, 2021; Lieder, Griffiths, Huys, & Goodman, 2018). We similarly use a resource-rationality framework to analyze adults intervention choices and judgments in our tasks. This involves first considering the impact of both information and complexity on inferential success, and second, modeling intervention selection as driven by a goal of maximizing the expected informativeness of the evidence while minimizing the expected inferential cost of processing the evidence.

### 1.5. Overview of experiments

We conducted two experiments to test how people actively learn causal structures in a continuous-time setting. In both experiments, we manipulated the reliability of the cause–effect delays and included a range of acyclic and cyclic causal structures. In Experiment 1, we only allowed participants to activate components while in Experiment 2, we also allowed them to block components.

In line with our normative account of causal inference in this setting, we hypothesized that performance would be lower in the irregular delay condition given that evidence about what caused what is more ambiguous. In line with our bounded inference account, we also hypothesized that performance would be worse in cyclic systems given the likely increase in event density, interdependence and concomitant complexity. However, we further expected accuracy to depend on the quality and reactivity of participants’ intervention choices. Thus, we also examine whether and how participants’ intervention selection differs across devices and delay conditions, asking to what extent intervention choice is reactive to the behavior of the device being explored, and whether this reactivity reflects rational anticipation and active management of expected information gain and evidential complexity.



**Fig. 3.** Experimental procedure. (a) Experimental interface. Up to 6 interventions could be performed by clicking on the components during the 45 s 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 needed 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). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

## 2. Experiment 1: Causal structure induction in continuous time

### 2.1. Methods

#### 2.1.1. Participants

Seventy-four participants (40 female, 34 male, aged  $30 \pm 11$ ) 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 *Incentives* section). Nine additional participants were tested but removed from the analysis because they left the default “unconnected” connection judgment for all causal component pairs for 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, & Lagnado, 2018; Coenen, Ruggeri, et al., 2019).<sup>1</sup>

#### 2.1.2. Design

Participants were randomly assigned to one of two delay conditions: (1) reliable delays ( $M \pm SD = 1.5 \pm 0.1$  s, i.e.  $\alpha = 200, \beta = \alpha/1500$  in the Gamma distribution) or (2) unreliable delays ( $1.5 \pm 0.7$  s, i.e.  $\alpha = 5, \beta = \alpha/1500$ , see Fig. 1c).

Participants were asked to investigate abstract causal “devices” connected by hidden causal links (Fig. 3a). The causal links produce point events in the form of activations of the device’s components over time. For causally related components, an activated component will probabilistically activate each of its effect components once after some delay. All causal connections 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). Participants were informed and tested on this in the instructions.

Each participant learned about 12 test devices with either 3 or 4 components, including 6 acyclic structures and 6 cyclic structures (Fig. 4). The acyclic structures were chosen to exemplify a variety of causal relationships including common effects (i.e. “colliders” Acyclic1 and Acyclic4), chains (Acyclic2 and Acyclic5), common causes (i.e. “forks”, Acyclic3 and Acyclic6). The cyclic devices were chosen so as to approximately match the number of edges in the acyclic systems while investigating a variety of arrangements. These included full loops (Cyclic1 and Cyclic4) and short loops with incoming connections (Cyclic2 and Cyclic5) and outgoing connections (Cyclic3 and Cyclic6).

**Interface.** Fig. 3 shows the task interface. For each device, participants saw the 3- or 4-components visualized as gray circles evenly spaced on a white background. Participants had 45 s to learn about how the components were connected. During this time, they could intervene and activate components by left-clicking on them up to 6 times. Intervened-on components were marked by a “+” symbol (Fig. 3a). All activated components turned yellow for 200 ms and then returned to gray (Fig. 3b). At the beginning, all components were inactive, and no connecting links 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, both directions, or no connection. Each click cycled through the options ( $A \rightarrow B, B \leftarrow A, A \leftrightarrow B, \text{no relationship}$ ) in a random order varied between participants. Participants confirmed their choices by clicking a confirm button that appeared in the middle (Fig. 3c). Links did not disappear after being confirmed, so participants were still able to update earlier judgments. What participants had marked at the end of 45 s was automatically registered as the final judgment for that trial. At the end of the

<sup>1</sup> The experimental procedure, data, and analysis code are available at: [https://github.com/tianweigong/time\\_and\\_intervention](https://github.com/tianweigong/time_and_intervention).



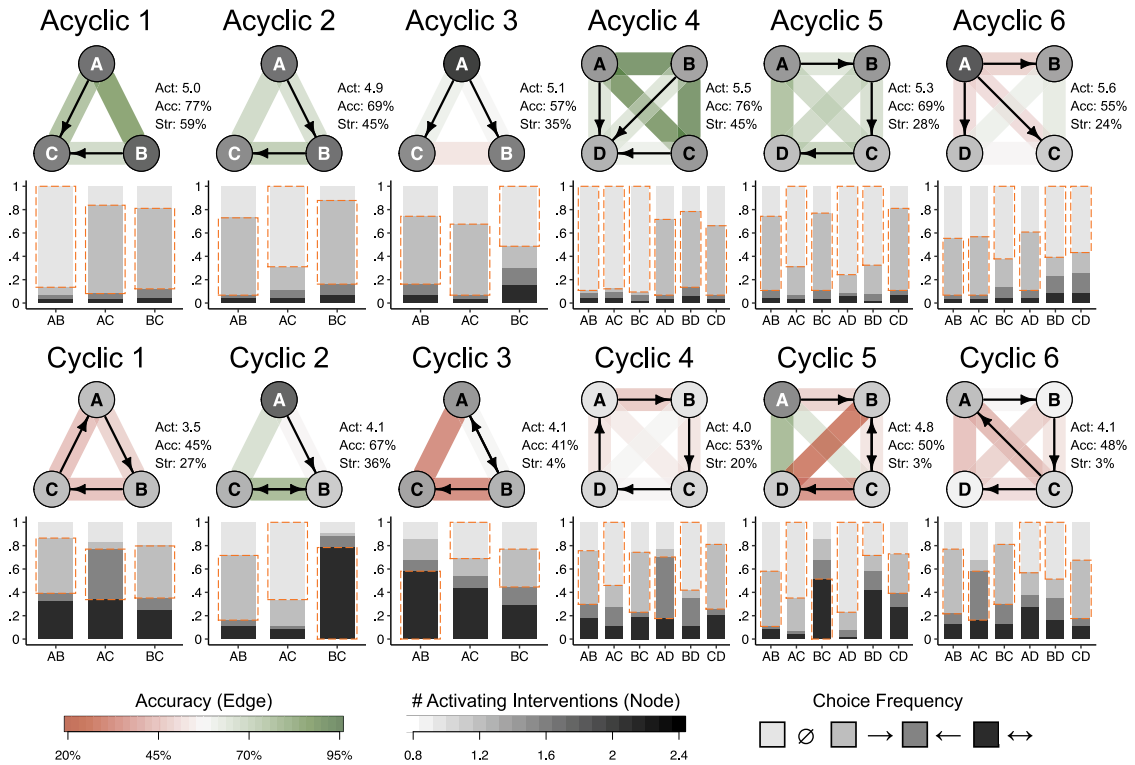


Fig. 4. Causal link identification and activating intervention choices in Experiment 1. Color edge shading indicates accuracy. Node shading indicates activating intervention choice prevalence by component. Bar plots show the proportion of different choices on each link (e.g. for AB, “∅” means “no connection between A and B”; “→” means “ $A \rightarrow B$ ”; “←” means “ $A \leftarrow B$ ”; “↔” means “ $A \leftrightarrow B$ ”) with orange used to highlight the ground truth. Note: Act = average number of activating interventions performed; Acc = mean accuracy; Str = proportion of participants who detected the whole structure correctly. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

trial, participants received feedback about which connections they had marked correctly or incorrectly (Fig. 3d). Since any pair of components might be unconnected, have a directed ( $A \rightarrow B$  or  $B \rightarrow A$ ), or bidirectional ( $A \leftrightarrow B$ ) causal connection, the response space includes 64 possible structures for 3-variable devices and 4096 possible structures for 4-variable devices of which exactly one truly reflects the hidden causal structure.

**Incentives.** We incentivized participants to mark the correct causal links as early as possible within the trial by rewarding them based on their accuracy at a random time point during each problem. This bonus scheme means it is in participants’ interest to register their best guess accurately and early, and to update it whenever their conclusions change during a learning episode (cf. Bramley, Dayan, et al., 2017). The scheme also shapes the nature of an ideal intervention strategy, meaning one should balance the benefits for intervention selection of waiting until one knows more, against the opportunity cost of waiting too long and missing out on what could have been learned from an earlier intervention. In order to perform well in our task, learners need not only consider *how* and *where* to intervene next, but also *when* to do so.

### 2.1.3. Procedure

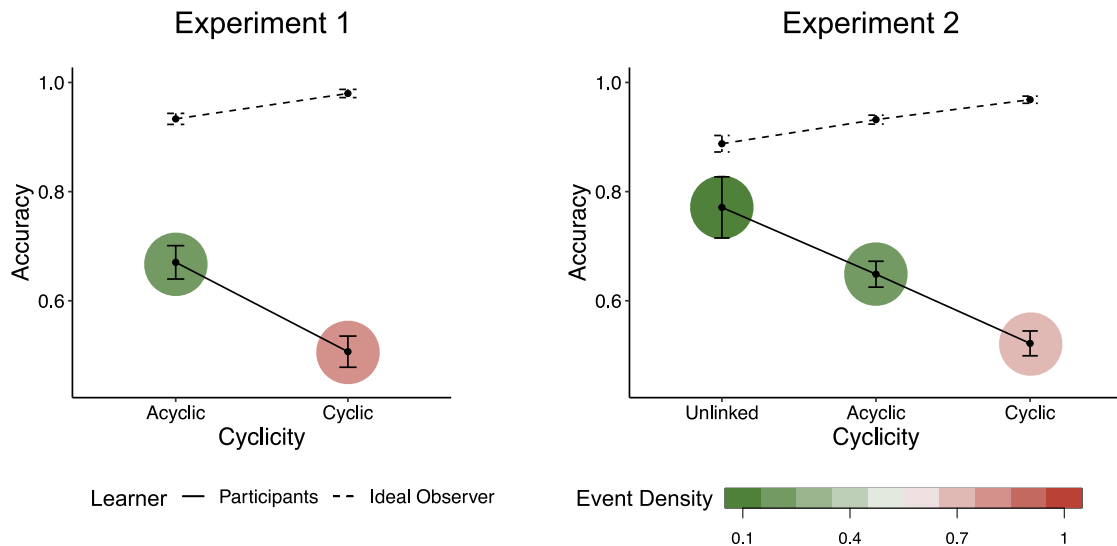
In the main task, each participant faced 12 test devices in random order with randomly positioned and unlabeled components. Prior to the inference task, participants completed instructions, a practice trial and comprehension checks. Participants were told that they would be investigating the causal structure of a number of abstract “devices”. In the instructions, participants were trained on the true cause-effect delays in their condition and shown a video example of a device with its causal links revealed. They were then trained on how to provide 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 Fig. 4).

## 2.2. Results

We first report participants’ judgment accuracy (i.e. what proportion of connections participants correctly identify at the end of each trial) by delay condition (reliable vs. unreliable), device type (acyclic vs. cyclic), and number of components (3 vs. 4). We then

**Table 1**  
Accuracy separated by conditions.

	Delay reliability		Structure cyclicity			Structure nodes	
	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%



**Fig. 5.** Participants vs. the ideal observer's accuracy and event density upon human generated evidence. Error bars indicate 95% confidence intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

discuss characteristic error patterns under specific causal structures. Our accuracy analyses use linear mixed-effect models (LMMs) including random slopes and intercepts for subject ID and structure type (Brauer & Curtin, 2018). For all LMMs we report standard coefficient estimates  $\beta$ s (that show how many units of standard deviations the outcome variable changes when the independent variable changes from one condition to the other),  $t$  values, significance, and 95% confidence intervals (CI). We then compare participants' trial-by-trial accuracy against the predictions of a normative inference model. We explore whether deviations from normative responding are related to the density of events in that trial as a basic index of complexity. We will then focus on participants' intervention choices and explore whether interventions are driven by a trade-off between the expected evidence strength and complexity.

**Accuracy.** Participants confirmed their causal judgments  $2.45 \pm 1.31$  times per trial. Final judgments — i.e. what participants had marked at the end of the trial — 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 their initial judgments — i.e. what participants had marked as their first answers — in the 79% of trials in which participants made more than one judgment (initial accuracy:  $58\% \pm 30\%$ ,  $\beta = 0.11$ ,  $t = 2.36$ ,  $p = .024$ ,  $CI = [0.02, 0.20]$ ). 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  $d = 1.86$ ) and were above chance for all 12 structures taken individually in both the reliable ( $t(35) > 4.15$ ,  $ps < .001$ ) and the unreliable condition ( $t(37) > 3.49$ ,  $ps < .01$ ) with the exception of `Cyclic1` (unreliable:  $t(37) = 1.94$ ,  $p = .06$ , Fig. 4). Table 1 shows accuracy separated by condition. There was a main effect of structure cyclicity ( $\beta = 0.50$ ,  $t = 2.69$ ,  $p = .026$ ,  $CI = [0.14, 0.86]$ ) such that the accuracy was higher for acyclic than cyclic structures. There was no main effect of delay reliability ( $\beta = 0.21$ ,  $t = 1.47$ ,  $p = .15$ ,  $CI = [-0.07, 0.49]$ ), or of number of components ( $\beta = 0.02$ ,  $t = 0.09$ ,  $p = .93$ ,  $CI = [-0.34, 0.37]$ ). Nor were there any two- or three-way interactions ( $ps > .10$ ).

**Error patterns.** Participants were best at inferring the structure of colliders (`Acyclic1` and `Acyclic4`, see Fig. 4). These structures were naturally simple in their evidence since no intervention would cause more than one effect. For the three-component chain (`Acyclic2`), 15% of participants added an erroneous additional direct connection  $A \rightarrow C$ . Similarly, for the four component chain (`Acyclic5`), participants frequently also added one or more “short cut” links from  $A \rightarrow C$  (12%),  $A \rightarrow D$  (8%) or  $B \rightarrow D$  (13%) in addition to the true connections. These errors cohere with previous findings suggesting that people rely on local computations when inferring causal structure, resulting in the addition of extraneous connections in chain structures (Davis et al., 2020; Fernbach & Sloman, 2009). No one mistook the chain `Acyclic2` for a fork with the same root component (i.e. `Acyclic3`), however 18%

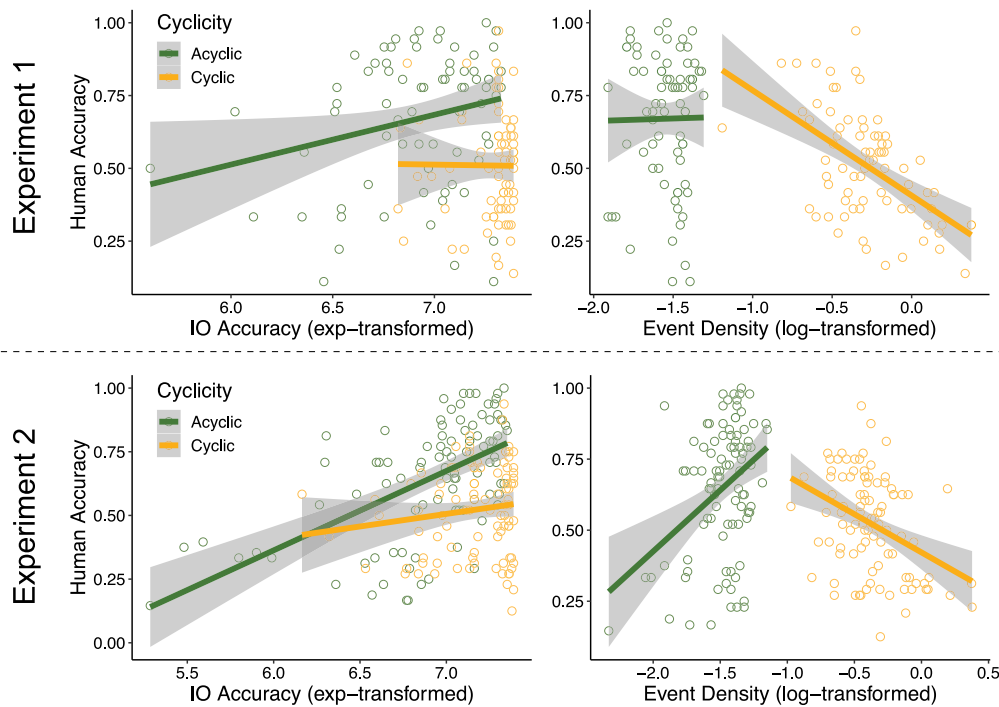


Fig. 6. Scatterplots of evidence informativeness (IO accuracy) and event density and participants' accuracy. Each data point shows an individual's average performance for acyclic or cyclic causal structures. One data point (4.96, 0.28) under the cyclic condition was removed from the upper left panel for visualization.

mistook the fork Acyclic3 for a chain with the same root component ( $A \rightarrow B \rightarrow C$  or  $A \rightarrow C \rightarrow B$ ). This lines up with the idea that people tend to fall-back on temporal order as a cue to causal order (Bramley, Gerstenberg, Mayrhofer, & Lagnado, 2018; McCormack et al., 2016), tending to link the effect components of a fork in whatever order they happened to activate. These error patterns did not differ significantly between reliable vs. unreliable groups ( $\chi^2$  tests,  $p_s > .10$ ).

For the cyclic structures, participants' judgments varied considerably so we focus on the individual-connection-level errors as shown in Fig. 4. In the full loops (Cyclic1 and Cyclic4), participants frequently judged directed or disconnected links as bidirectionally connected (Fig. 4, black bars). This was more prevalent in the unreliable group than in the reliable group (Cyclic1: 37% vs. 23%,  $\chi^2(1) = 4.31$ ,  $p = .04$ ; Cyclic4: 22% vs. 10%,  $\chi^2(1) = 11.41$ ,  $p < .001$ ), suggesting that reliable delays make it easier to detect full loop structures. This makes sense since regular delays produce much more sequentially reliable and predictable patterns of reactivation. Participants had relatively little trouble at identifying loops with incoming connections (Cyclic2). However, performance was very poor for structures comprised of feedback loops with outgoing connections (Cyclic3, Cyclic5, and Cyclic6). The outgoing component ( $C$  in Cyclic3,  $D$  in Cyclic5 and Cyclic6) was frequently taken to be a constituent of the feedback loop, often being assigned a bidirectional connection with one of the loop constituents. This is reasonable since, for these structures, recurrent and close-in-time events occurred not only at the components forming the loop themselves but also for the output components, making it difficult to tell which components were involved in actively sustaining the looping pattern of activations. Participants often connected an output component to the loop element that typically activated in close temporal proximity. For example, many participants marked a  $C \leftrightarrow A$  connection in Cyclic3,  $D \leftrightarrow B$  in Cyclic5, and  $D \leftrightarrow A$  in Cyclic6, in spite of the fact that this temporal proximity is really due to them sharing a common cause.

Participants were normally correct about whether the structure was cyclic or acyclic. Participants' structure judgments belonged to the correct class  $82\% \pm 38\%$  of time for the acyclic class and  $77\% \pm 42\%$  of time for the cyclic class. There was no difference in the frequency of mistaking cyclic for acyclic vs. acyclic for cyclic ( $t(73) = 1.15$ ,  $p = 0.25$ ), and the rate did not differ between reliable and unreliable delay conditions ( $p_s > .10$ ).

**Informativeness and event density.** We calculated the accuracy of an 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 was (cf. Bramley et al., 2015). The IO was more accurate in the reliable ( $97\% \pm 7\%$ ) than the unreliable ( $94\% \pm 12\%$ ) condition ( $\beta = 0.23$ ,  $t = 2.61$ ,  $p = .011$ ,  $CI = [0.06, 0.40]$ ). In contrast to human learners, the IO was more accurate in identifying the structure of cyclic structures ( $98\% \pm 8\%$ ) compared to acyclic ( $93\% \pm 11\%$ ) structures ( $\beta = 0.45$ ,  $t = 3.16$ ,  $p = .010$ ,  $CI = [0.17, 0.73]$ ), showing

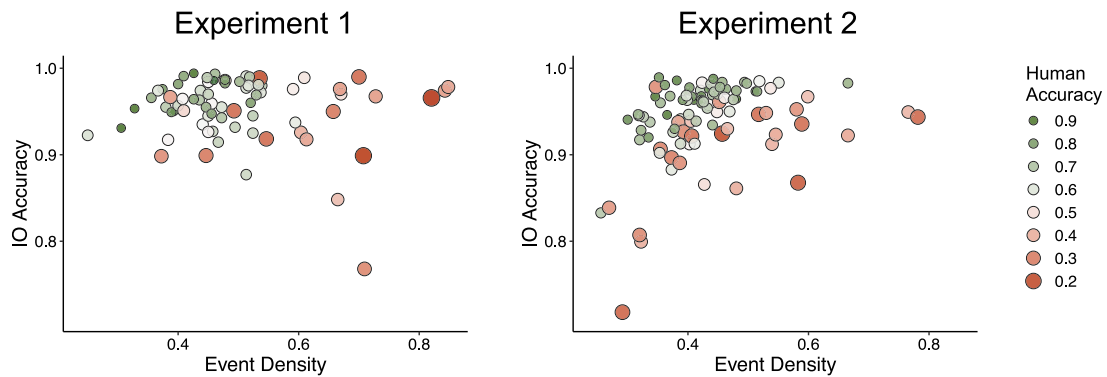


Fig. 7. Scatterplots of average final IO accuracy (indexing evidence informativeness) and event density (indexing evidence complexity) for each participant with color and size indicating that participants' judgment accuracy. Participants with higher accuracy generated evidence that was both more informative and less complex (the upper left area). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

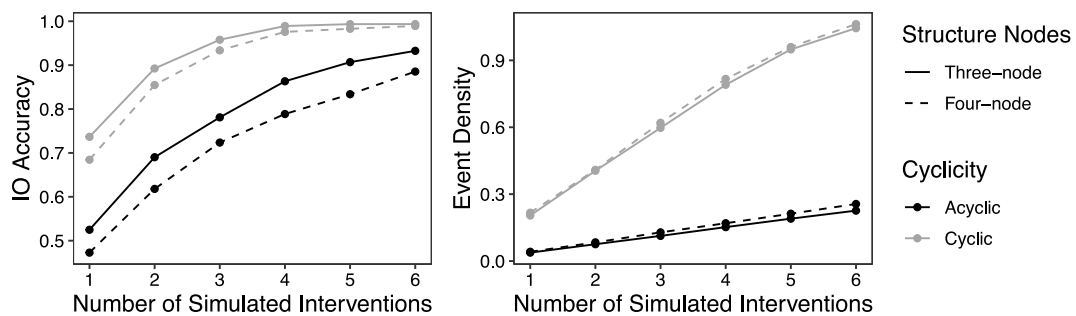


Fig. 8. Relationship between intervention count, evidence strength (IO accuracy) and complexity (event density) in simulated causal interactions. Simulations based on randomly activating a component in a random causal system (from Fig. 4) at  $t \in \{0, 7.5, 15, 22.5, 30, 37.5\}$  seconds so a full set of six interventions would be distributed evenly across 45 s.

the reverse pattern to human learners.<sup>2</sup> This suggests that, in principle, the cyclic devices produced more information about their underlying causal structure than the acyclic devices. IO accuracy in general does not predict human trial-by-trial accuracy ( $\beta = 0.05$ ,  $t = 1.40$ ,  $p = .178$ ,  $CI = [-0.02, 0.12]$ ). As Fig. 6 shows, the relationship between IO and participant accuracy was moderated by cyclicity with a positive association across acyclic structures ( $\beta = 0.16$ ,  $t = 3.61$ ,  $p = .002$ ,  $CI = [0.07, 0.24]$ ) but none for cyclic structures ( $\beta = -0.04$ ,  $t = 0.84$ ,  $p = .407$ ,  $CI = [-0.13, 0.05]$ ).

We use event density (number of events per second) as a basic index for how *complex* the generated evidence was. Event density differed dramatically between acyclic ( $0.22 \pm 0.06$ ) and cyclic ( $0.79 \pm 0.41$ ) devices ( $\beta = 1.39$ ,  $t = 11.80$ ,  $p < .001$ ,  $CI = [1.16, 1.63]$ ) and participants' accuracy was generally negatively related to density of events ( $\beta = -0.27$ ,  $t = 3.71$ ,  $p < .001$ ,  $CI = [-0.41, -0.14]$ ). However, as shown in Fig. 6, event density was negatively associated with accuracy on cyclic ( $\beta = -0.31$ ,  $t = 3.30$ ,  $p = .02$ ,  $CI = [-0.49, -0.12]$ ), but not acyclic ( $\beta = -0.01$ ,  $t = 0.25$ ,  $p = .810$ ,  $CI = [-0.11, 0.08]$ ) devices.

These results suggest that evidence complexity is critical in this task. For the IO, complexity is generally positively correlated with success. The more activations there are, the more information an ideal observer can use to reduce its uncertainty. However, non-ideal human learners clearly struggled to deal with complex evidence. As shown in Fig. 7, the best-performing participants were generally those who were able to generate evidence that would have enabled the IO to be highly accurate, but that was also low in event density. We later compare computational models that model the influence of complexity on human judgments and intervention selection, capturing the qualitative differences between cyclic and acyclic cases (see the section on *Modeling the Judgments*).

**When to intervene.** We now assess whether participants' intervention choices are qualitatively consistent with the idea that they choose interventions that generate strong evidence while minimizing evidential complexity. For example, participants may choose to perform fewer interventions when experiencing a large numbers of events, as tends to occur in cyclic structures and, to a lesser extent with structures with four components (Fig. 8). As shown in Fig. 9, on average, participants performed  $4.68 \pm 1.46$  out of the maximum of 6 interventions on each trial, performing about the same number in the unreliable ( $4.76 \pm 1.41$ ) and reliable

<sup>2</sup> 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 simply omit these trials from current analyses.

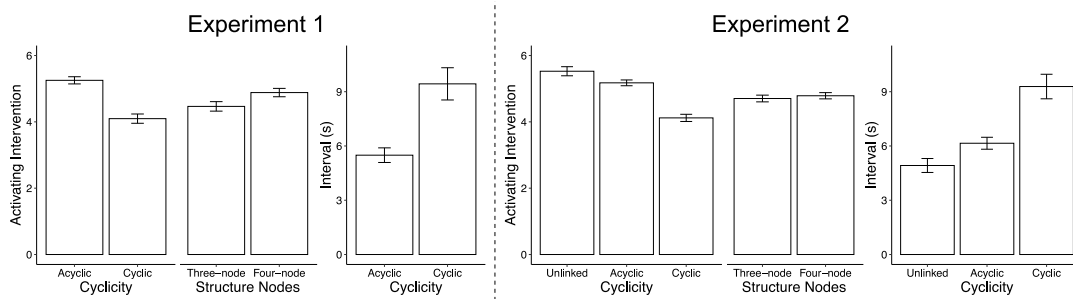


Fig. 9. The average numbers of activating interventions used and average time intervals between activating interventions under different structures. Error bars indicate 95% confidence intervals. Each data point shows an individual's average intervals for acyclic or cyclic causal structures.

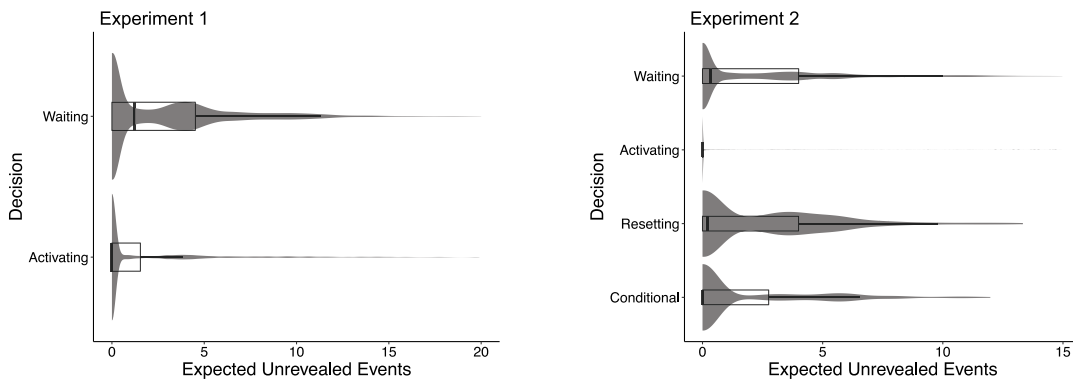


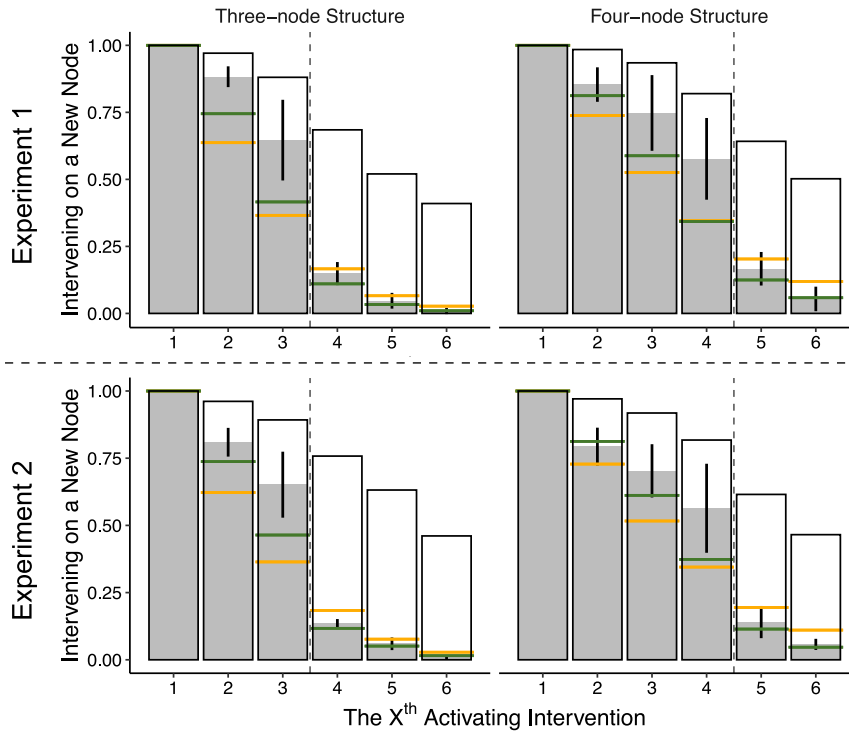
Fig. 10. Number of expected unrevealed events across all 1-second decision windows in all trials, as a function of whether an intervention is performed in that window. Windows for which participants performed more than one intervention were excluded. The densities are scaled for each experiment to have 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).

condition ( $4.58 \pm 1.51$ ,  $\beta = 0.12$ ,  $t = 0.93$ ,  $p = .36$ ,  $CI = [-0.14, 0.39]$ ). However, participants performed fewer interventions on cyclic ( $4.10 \pm 1.50$ ) than acyclic devices ( $5.25 \pm 1.16$ ,  $\beta = 0.79$ ,  $t = 6.24$ ,  $p < .001$ ,  $CI = [0.54, 1.04]$ ) and fewer on three-component ( $4.47 \pm 1.54$ ) than four-component devices ( $4.88 \pm 1.35$ ,  $\beta = 0.28$ ,  $t = 2.27$ ,  $p = .045$ ,  $CI = [0.04, 0.53]$ ). These results correspond to the simulation results in Fig. 8: In cyclic structures, even a few interventions allowed for ceiling level accuracy in principle, while the event density compounded dramatically with each additional intervention. This means that, in cyclic structures, the computational cost of additional interventions quickly outweighs the value of the new information. Event density also increased going from three- to four-component devices (at least for the structures we tested), but this increase comes alongside an increase in the amount of structure to be learned. This means that the evidence-strength gap between three- and four-component devices could be narrowed by intervening more frequently in four-component devices compared to three-component devices.

The average interval between each intervention depended on cyclicity ( $\beta = 0.67$ ,  $t = 7.22$ ,  $p < .001$ ,  $CI = [0.49, 0.86]$ ), with participants waiting longer before the next intervention when the structure being learned was cyclic ( $9.38 \pm 5.94$  s) rather than acyclic ( $5.49 \pm 2.64$  s, Fig. 9). There was no evidence for a difference in this measure between the unreliable and reliable delay conditions ( $7.55 \pm 5.10$  s vs.  $7.26 \pm 4.83$  s) or between three- and four-node problems ( $7.19 \pm 4.92$  s vs.  $7.63 \pm 5.01$  s). As shown in the example in Fig. 2b, even if the total number of events is identical, learning is easier when the events are more spread out across the trial. Intervention-spreading may be particularly important for cyclic structures where the event density is higher.

We test whether participants' tendency to wait longer under cyclic structures is driven by an anticipation of complexity. As a first pass, we calculated a moment-by-moment expectation of the level of computational cost of the upcoming evidence assuming no further intervention is performed. For this, we calculated the number of events expected to occur in the near future as a result of earlier activity (see *Modeling the Interventions* for more details). We can compare the moments in which participants did nothing with those in which participants performed an intervention. As shown in Fig. 10, 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 activating interventions yet). Yet in the 12% of time windows where they *did* intervene, the number of already-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:  $\chi^2(1) = 854.35$ ,  $p < 0.001$ ). This result suggests that participants tended to wait to perform their next intervention once there was not too much expected activity.

*Where to intervene.* An efficient sequence of interventions 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 activating components that showed



**Fig. 11.** Participants’ tendency to activate a node they have not intervened on previously as a function of intervention index. Black frames indicate the proportion of trials in which the participant performed at least this many interventions. Error bars indicate 95% confidence intervals. Yellow lines indicate performance 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 indicate the boundary after which the learner has performed enough interventions to have tried every component once. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

promise in producing effects. This repetition allows a learner to gather evidence about the order and the delay with which effects propagate through the system. This information is crucial for distinguishing between devices with overlapping causal structure. We see clear qualitative evidence of such exploration and exploitation in participants’ choices. Fig. 4’s node shading shows the aggregate proportion of interventions on each node in each structure. Participants’ interventions were relatively evenly distributed across components for most devices. They had a slight tendency to activate 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. Fig. 11 shows how frequently participants selected a novel component to intervene on as a function of serial intervention position within the trial. This is shown in Fig. 11, and we compared participants’ performance against chance (i.e. the random intervener) as well as against the choices of an idealized expected-information-gain maximizing intervener (i.e. the EIG intervener, see the section on *Modeling the Interventions* for more details) 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 ( $t(73) > 10.91, p < .001$ ). This shows that participants were not intervening randomly and suggests that they typically began by exploring system components they had not activated yet. The simulated informationally efficient intervener shows a similar pattern the first several interventions (Fig. 11). The efficient intervener’s decisions, 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.

### 2.3. Discussion

In Experiment 1, we showed that people are able to infer causal structure through active intervention in a challenging continuous-time learning setting. We found that participants, had different error patterns to an ideal observer model, in particular making more accurate judgments about acyclic structures than cyclic structures while the ideal observer had the reverse pattern. We

also found that differences in accuracy across conditions were associated with differences in the character of the evidence. The informativeness of evidence predicted participants' performance in acyclic structures, but the complexity of evidence appeared to dominate participants' performance in cyclic structures where it was generally higher. Participants who were able to generate evidence that was both informative but not overly complex tended to perform best overall. We take this to support our central idea that managing computational cost plays an important role in interventional decisions and success in the real-time causal learning setting.

Intervention choices were partly shaped by a drive to control computational demands. In terms of when to intervene, participants performed fewer interventions and waited longer between them on cyclic structures that tended to produce more events. They also tended to perform more interventions on four-node structures yielding a similar number of events as for three-node structures but presumably responding to the greater initial uncertainty (larger space of structure possibilities). When the expected upcoming evidential complexity was already high, participants were more likely to wait rather than activate another component to produce more events.

In terms of which components participants would target, we found they used their interventions to systematically explore the devices, tending to select a hitherto untested component for their first few interventions, qualitatively in line with the behavior of an efficient information maximizing agent. Participants also showed a tendency to repeat-intervene on causally "central" components once these were 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 previously been framed as a heuristic Positive Testing Strategy (Coenen et al., 2015; Steyvers et al., 2003) on the grounds that it is often correlated with expected information yet much easier to calculate. Positive testing can be very poor in the atemporal setting because multiple causal influences from the root component overshadow one another since all effects are revealed at once. However, in the continuous-time point-event case, intervening on a suspected root will often generate rich and diagnostic evidence through the delays and order variability in the propagation of the activity through the system (Bramley, Gerstenberg, Mayrhofer, & Lagnado, 2018). We will test the extent to which participants' specific where-to-intervene choices reflect an information gain norm in our model fitting to follow (see the section on *Modeling the Interventions*).

### 3. Experiment 2: Activating and blocking

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. We will examine whether the relationship between ideal observer accuracy, human accuracy and event density 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 reduce the complexity of evidence without substantially reducing its diagnosticity about the causal relationships.

#### 3.1. Methods

##### 3.1.1. Participants

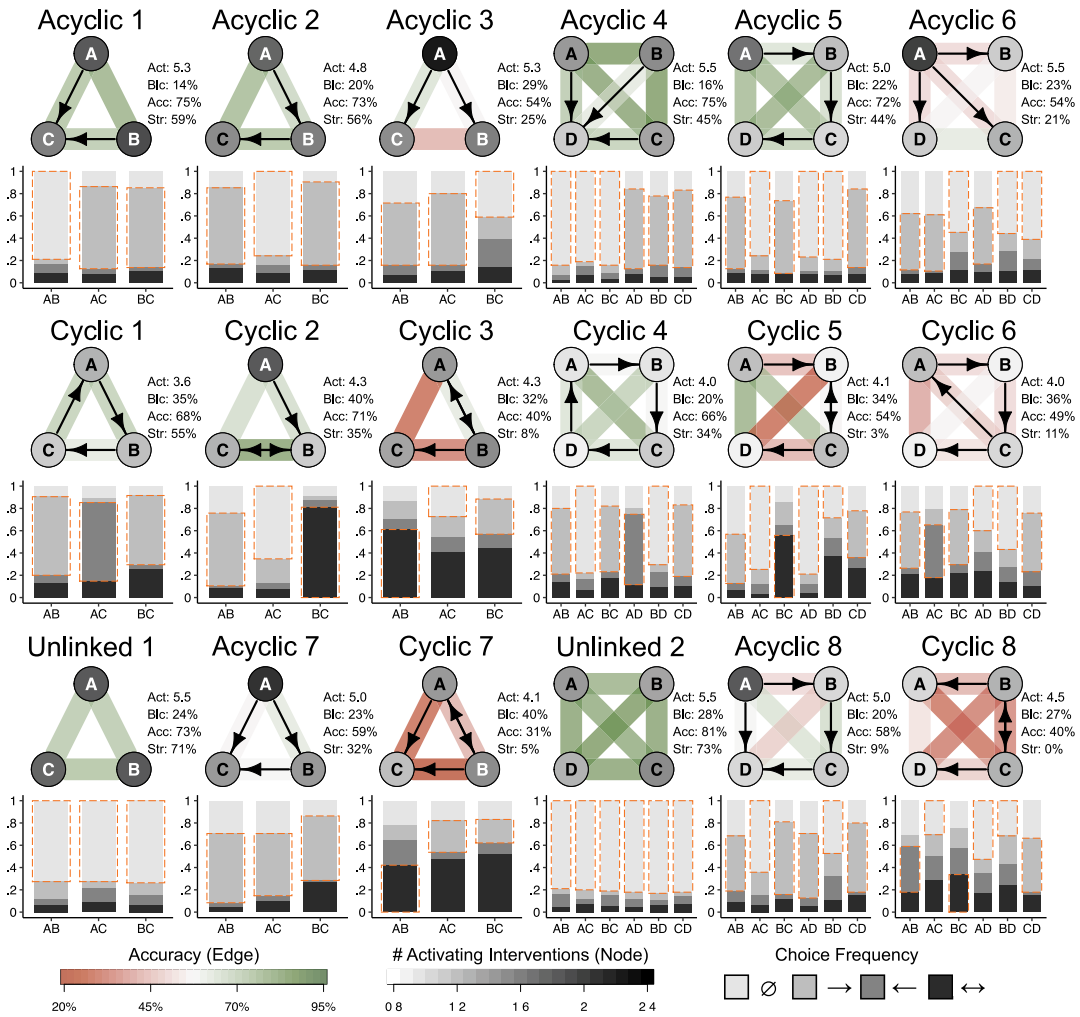
95 participants (54 female, 40 male, 1 nonbinary, aged  $36 \pm 12$ ) were recruited from Prolific Academic and were randomly assigned to reliable-delay ( $N = 48$ ) or unreliable-delay ( $N = 47$ ) condition. They received a basic payment of £1 and a bonus depending on performance as in Experiment 1. Fourteen additional participants were tested but removed from the analysis because they reported 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 ( $N = 7$ ).

##### 3.1.2. Design & procedure

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 (Fig. 3b). 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. While participants were limited to 6 activations (as in 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, Fig. 12), two densely connected cyclic structures (Cyclic7 and Cyclic8), as well as two devices with no causal connections 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 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 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



**Fig. 12.** Causal link identification and activating intervention choices in Experiment 2. Color edge shading indicates accuracy. Node shading indicates activating intervention choice prevalence by component. Bar plots show the proportion of different choices on each link (e.g. for AB, “∅” means “no connection between A and B”; “→” means “A → B”; “←” means “A ← B”; “↔” means “A ↔ B”) with orange used to highlight the ground truth. Note: Act = average number of activating interventions performed; Blc = proportion of participants who used blocking; Acc = mean accuracy; Str = proportion of participants who detected the whole structure correctly. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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 make providing judgments more seamless, 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 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.

### 3.2. Results

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

**Accuracy.** Participants registered judgments  $4.39 \pm 2.43$  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.



**Table 1** shows participants' accuracy separated by conditions. 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  $d = 1.99$ ). The average accuracy for all 18 structures were above chance in the reliable condition ( $t(47) > 4.01$ ,  $ps < .001$ ) and the unreliable condition ( $t(46) > 3.39$ ,  $ps < .01$ ) with the exception of *Cyclic7* (reliable:  $t(47) = 2.01$ ,  $p = .05$ ; unreliable:  $t(46) = 0.56$ ,  $p = .58$ , **Fig. 12**).

Unlike in Experiment 1, participants' performance only differed between unlinked and cyclic structures ( $\beta = 0.72$ ,  $t = 2.26$ ,  $p = .04$ ,  $CI = [0.10, 1.35]$ , **Fig. 5**), with a marginally significant difference between acyclic and cyclic structures ( $\beta = 0.37$ ,  $t = 1.85$ ,  $p = .09$ ,  $CI = [-0.02, 0.75]$ ). There was no main effect of delay reliability ( $\beta = 0.03$ ,  $t = 0.20$ ,  $p = .84$ ,  $CI = [-0.24, 0.30]$ ) or the number of components ( $\beta = 0.08$ ,  $t = 0.34$ ,  $p = .74$ ,  $CI = [-0.37, 0.53]$ ), nor were there any two- or three-way interactions ( $ps > .10$ ).

**Error patterns.** **Fig. 12** shows the types of errors people made in inferring causal structures. For chain structures (*Acyclic2*, *Acyclic5*), there were no systematic errors in mistaking them as fully-connected structures, or fork structures (less than 5%). Similar to Experiment 1, 15% of participants mistook the fork structure *Acyclic3* as a chain structure ( $A \rightarrow B \rightarrow C$  or  $A \rightarrow C \rightarrow B$ ), while 12% mistook it for a fully-connected structure by adding a directed link between two child nodes. In the case of the fully-connected structure *Acyclic7*, 15% of participants disregarded the link between  $A \rightarrow C$ , while 10% confused it for a cyclic structure  $A \rightarrow B \leftrightarrow C$ . For *Acyclic8*, 13% of participants confused  $B \rightarrow C$  with  $A \rightarrow C$ , or  $C \rightarrow D$  as  $B \rightarrow D$ . The error patterns did not significantly differ between unreliable and reliable delay conditions ( $\chi^2$  tests,  $ps > .10$ ). Similar to Experiment 1, these error pattern results seem consistent with the idea that reliance on local computation and simple event order played a role in some participants' judgments (**Bramley, Gerstenberg, Mayrhofer, & Lagnado, 2018; Burns & McCormack, 2009; McCormack et al., 2016**).

For cyclic structures, participants in Experiment 2 also performed poorly for structures that contained one or more output components. The output components were frequently judged to be constituents of the feedback loop. Participants tended to connect components whose activations often occurred close in time. This pattern was replicated in the two new structures *Cyclic7*, *Cyclic8*. For instance, participants frequently marked erroneous bidirectional  $A \leftrightarrow C$  and  $B \leftrightarrow C$  connections in *Cyclic7* and  $A \leftrightarrow C$  and  $B \leftrightarrow D$  in *Cyclic8*.

As with Experiment 1, participants could generally tell whether a structure was cyclic or acyclic regardless of whether they got all causal connections correct. They choose the correct class  $70\% \pm 46\%$  of time for acyclic structures (excluding the unlinked structures) and  $80\% \pm 40\%$  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 event density.** Based on observing participants' interventions, the ideal observer performed significantly better at identifying the structure of cyclic ( $97\% \pm 9\%$ ) than unlinked ( $89\% \pm 11\%$ ) devices ( $\beta = 0.74$ ,  $t = 3.59$ ,  $p = .004$ ,  $CI = [0.33, 1.14]$ , **Fig. 5**), but not acyclic ( $93\% \pm 12\%$ ) devices ( $\beta = 0.37$ ,  $t = 1.85$ ,  $p = .09$ ,  $CI = [-0.02, 0.75]$ ). It also performed better for three-node ( $96\% \pm 10\%$ ) than four-node ( $93\% \pm 11\%$ ) devices ( $\beta = 0.40$ ,  $t = 2.62$ ,  $p = .02$ ,  $CI = [0.10, 0.69]$ ). Unlike in 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 ( $\beta = 0.10$ ,  $t = 3.28$ ,  $p = .003$ ,  $CI = [0.04, 0.16]$ ), with a positive relationship in acyclic devices ( $\beta = 0.15$ ,  $t = 3.27$ ,  $p = .008$ ,  $CI = [0.06, 0.24]$ ) but no relationship for cyclic devices ( $\beta = 0.04$ ,  $t = 0.12$ ,  $p = .22$ ,  $CI = [-0.03, 0.11]$ , **Fig. 6**).

Event density differed between acyclic ( $0.23 \pm 0.08$ ) and cyclic ( $0.71 \pm 0.39$ ) trials ( $\beta = 1.31$ ,  $t = 11.86$ ,  $p < .001$ ,  $CI = [1.09, 1.58]$ ), and between unlinked ( $0.12 \pm 0.02$ ) and cyclic trials ( $\beta = 1.61$ ,  $t = 9.71$ ,  $p < .001$ ,  $CI = [1.28, 1.93]$ ). Participants' accuracy was generally negatively associated with event density ( $\beta = -0.20$ ,  $t = -6.15$ ,  $p < .001$ ,  $CI = [-0.26, -0.14]$ ). However, this also depended on cyclicity. Event density was negatively correlated with human accuracy in cyclic trials ( $\beta = -0.21$ ,  $t = 3.88$ ,  $p = .006$ ,  $CI = [-0.32, -0.11]$ ), but there was no significant relationship for acyclic trials ( $\beta = 0.10$ ,  $t = 1.73$ ,  $p = .12$ ,  $CI = [-0.01, 0.22]$ , **Fig. 6**). As shown in **Fig. 7**, participants who generated evidence with both high informativeness and low complexity performed better in general.

**When to activate.** Participants performed  $4.75 \pm 1.46$  activations on average on each trial. This did not differ across reliable ( $4.75 \pm 1.42$ ) and unreliable ( $4.74 \pm 1.49$ ) conditions or across three-node ( $4.70 \pm 1.51$ ) and four-node ( $4.79 \pm 1.40$ ) structures. However, participants performed fewer activations on cyclic ( $4.12 \pm 1.52$ ) compared to acyclic devices ( $5.18 \pm 1.23$ ,  $\beta = 0.72$ ,  $t = 6.97$ ,  $p < .001$ ,  $CI = [0.52, 0.93]$ ), or compared to unlinked devices ( $5.53 \pm 0.96$ ,  $\beta = 0.97$ ,  $t = 6.16$ ,  $p < .001$ ,  $CI = [0.66, 1.27]$ , **Fig. 9**). Participants also waited longer to perform activations in cyclic than acyclic structures ( $9.26 \pm 5.57$  s vs.  $6.15 \pm 2.54$  s,  $\beta = 0.53$ ,  $t = 7.63$ ,  $p < .001$ ,  $CI = [0.39, 0.66]$ ) or unlinked structures ( $4.91 \pm 2.33$  s,  $\beta = 0.79$ ,  $t = 7.57$ ,  $p < .001$ ,  $CI = [0.58, 0.99]$ ), and longer in acyclic than unlinked structures ( $\beta = 0.26$ ,  $t = 2.74$ ,  $p = .02$ ,  $CI = [0.07, 0.44]$ , **Fig. 9**). There were, again, more expected unrevealed events on seconds where participants 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:  $\chi^2(1) = 2423.60$ ,  $p < 0.001$ , **Fig. 10**).

**When to block.** Participants performed blocks on 27% of trials (949 times in 1710 trials,  $0.55 \pm 1.17$  per trial). 75 of 95 participants (79%) used blocking at least once. Given that the frequency of blocking was much sparser than activations, we coded trials as 1 (used blocks) or 0 (no blocks) in our statistical analyses and fitted logistic regression models to explore when blocking was used.

We found that the propensity to block differed neither between the reliable (28%) and unreliable (26%) delay conditions ( $\beta = 0.12$ ,  $z = 0.87$ ,  $p = .386$ ,  $CI = [-0.15, 0.38]$ ), nor between three-node (29%) and four-node (25%) devices ( $\beta = 0.07$ ,  $z = 0.50$ ,  $p = .620$ ,  $CI = [-0.20, 0.33]$ ). However, participants were more likely to use blocks in cyclic (33%) than acyclic (21%) structures ( $\beta = 0.61$ ,  $z = 5.16$ ,  $p < .001$ ,  $CI = [0.37, 0.84]$ ). Surprisingly, propensity to use blocks when facing unlinked (26%) structures did not differ significantly from cyclic or acyclic structures ( $ps > .05$ ). We had anticipated participants would be less likely to block in

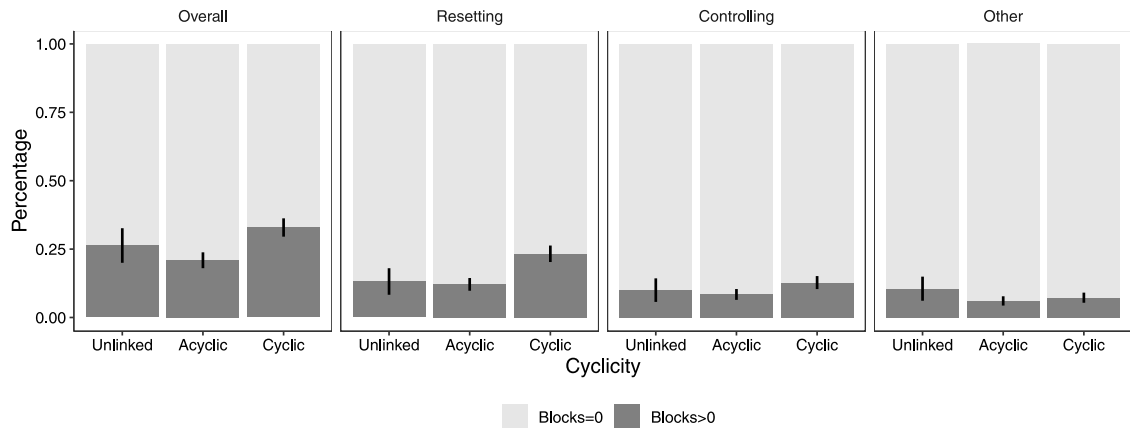


Fig. 13. Percentages of blocking behaviors. Error bars indicate 95% confidence intervals.

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 empirical signatures: (1) Blocking in combination with activating to control for confounding causal paths and (2) blocking to reset the device.

For both of these uses, we derived simple operationalizations. We take “Controlling” blocks to be those 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 the blocked one(s). This way, sub-networks can be investigated through an activation without the possibility of interference from any pathways through the blocked component. In contrast, the “Resetting” category includes 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. Both forms of blocking reduce density of events experienced during the trial but do so in conceptually different ways (See examples in Fig. 2b). “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 acyclic (12%,  $\beta = 0.79$ ,  $z = 5.53$ ,  $p < .001$ ,  $CI = [0.51, 1.07]$ ) or unlinked (13%,  $\beta = 0.69$ ,  $z = 2.96$ ,  $p = .003$ ,  $CI = [0.25, 1.17]$ ) devices. There was no difference between acyclic and unlinked structures ( $\beta = 0.09$ ,  $z = 0.38$ ,  $p = .704$ ,  $CI = [-0.41, 0.56]$ ). Similarly, participants performed more Controlling blocks in cyclic (13%) than acyclic (8%,  $\beta = 0.47$ ,  $z = 2.72$ ,  $p = .007$ ,  $CI = [0.13, 0.82]$ ) 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%,  $\beta = 0.64$ ,  $z = 2.24$ ,  $p = .025$ ,  $CI = [0.06, 1.19]$ ) but not cyclic structures (7%). In sum, both Resetting and Controlling were used more on cyclic devices (see Fig. 13).

We checked whether performing blocks is related to participants’ accuracy. We compared accuracy of different structures in Experiment 1 and 2 (Fig. 4 vs. Fig. 12) and found that the accuracy of full loops *Cyclic1* ( $t(167) = 3.62$ ,  $p < .001$ ) and *Cyclic4* ( $t(167) = 2.53$ ,  $p = .012$ ) significantly differed between two experiments. We checked whether the blocking function accounted for the difference. For *Cyclic1*, participants in Experiment 2 performed at least one Resetting (26%) were more accurate than those who did not make this kind of block ( $t(93) = 2.42$ ,  $p = .018$ , Cohen’s  $d = 0.56$ ). However, Controlling blocks (17%) did not make a significant difference ( $t(93) = 1.48$ ,  $p = .141$ ). This finding was replicated in *Cyclic4* that accuracy was positively associated with the use of Resetting blocks (16%,  $t(93) = 3.58$ ,  $p < .001$ , Cohen’s  $d = 1.01$ ), but not with Controlling blocks (6%,  $t(93) = 1.87$ ,  $p = .06$ ). This indicated that Resetting blocks may be more helpful than Controlling blocks. Given that the performance was better for several cyclic structures, there was some benefit to having the blocking ability.

**Where to activate.** Similar to Experiment 1, participants tended to explore the devices initially by activating untested components. Participants were more likely than chance to activate an untested component with their second and third activating interventions (and their fourth for four node devices,  $p_s < .001$ ), which is in line with how the informationally efficient intervener behaves (Fig. 11).

**Discussion.** In Experiment 2, we allowed participants to use blocking as a tool for 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 system but also to selectively inhibit it. As found in Experiment 1, evidence informativeness positively predicted participants' performance in acyclic structures while evidential complexity negatively predicted performance in cyclic structures. Accuracy was less strongly associated with structure cyclicality than in Experiment 1, which 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 the finding that people performed fewer activations and waited longer to perform their next activation in cyclic structures where expected computational cost was generally higher.

Participants used blocking in only a quarter of trials. However, when blocking was employed it was used in sensible ways that primarily managed inferential complexity. Participants 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 management of computational cost into consideration when choosing how to intervene in real time. We categorized blocks according to two potential goals: Resetting the system and Controlled testing — combining a block with an activation to test a subsystem in isolation. Both of these strategies were more likely to be employed after moments of high expected complexity.

#### 4. Modeling the judgments

The following two sections detail our quantitative analysis of the role of complexity in shaping participants' causal judgments and intervention choices. We compare a set of computational models to demonstrate that: (1) Participants' causal judgments were affected by evidential informativeness and complexity and (2) participants' interventions strike a balance between the informativeness and complexity of future evidence. Readers less interested in technical detail can safely skip ahead to the General Discussion.

To quantitatively test the idea that evidence complexity is not just positively related to informativeness, but also impacts human performance directly, we built a computational-level model that assumes human causal judgments  $q \in \mathbf{Q}\{X \rightarrow Y, X \leftarrow Y, X \leftrightarrow Y, X \emptyset Y\}$  are a noisy version of the ideal observer's posterior marginalized across connections  $\text{IO}_q$ , where the noise degree depends on the density, and hence complexity, of the evidence. We capture this with a dynamic softmax function (Luce, 1959):

$$P(\text{judgment} = q) = \frac{\exp(\text{IO}_q / (\tau_1 N + \tau_2))}{\sum_{q' \in \mathbf{Q}} \exp(\text{IO}_{q'} / (\tau_1 N + \tau_2))} \quad (1)$$

where  $N$  denotes a trial's event density (average number of events per second). The judgment temperature component is thus a linear function of events  $f(N) = \tau_1 N + \tau_2$  with two parameters  $\tau_1, \tau_2 \in (0, +\infty)$  that are constant across trials, while  $N$  varies across trials depending on what interventions are performed and how the system reacts to them. As  $f(N) \rightarrow +\infty$ , model predictions become more uniform, while if  $f(N) \rightarrow 0$  the predictions become more deterministic and increasingly resemble those of an ideal observer maximizing over the posterior.

We fit this model with participants' choices across two experiments and compare it to a baseline model that made random judgments, and a informativeness-based model that only considers IO judgments by omitting  $\tau_1 N$  from Eq. (1). We used 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, Navarro, Perfors, & Steyvers, 2017). Our cross-validation scheme is conservative, since it requires a unified explanation for human data despite different causal devices exhibiting markedly different characteristic dynamics.

Table 2 shows the results. In both Experiment 1 and 2, the model that combines informativeness and complexity outperforms the informativeness-only model. Individual results are relatively similar between the two models in Experiment 1, but more-strongly favor the combination model in Experiment 2 where there were more data points for a single person (12 vs. 18 data points in Experiment 1 vs. Experiment 2). These results suggest that, while more complex evidence carries more information on average, its complexity takes a toll on human performance, presumably due to our cognitive limitations.

#### 5. Modeling the interventions

The final part of our analysis describes a computational account of complexity-sensitive intervention selection and compares it to participants' intervention choices in both experiments. Intervention selection is the problem of choosing what to do now, in order to support future learning. Normatively, this depends on the learner's prior over causal structures  $P(S)$  at the point of the decision, which in turn depends on the already-observed data  $\mathbf{d}$ , and earlier interventions  $\mathbf{i}$ . In our first experiment, where participants can activate but not block, each participant must choose, at each moment in time, between intervening on one of the components or doing nothing, leading to the intervention space  $\mathbf{I} = \{a_A, a_B, a_C, \emptyset\}$  for three node systems. If the learner has used all their activation chances, this reduces to just the option of doing nothing  $\emptyset$ . In our second experiment, where participants and models could 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.  $\mathbf{I} = \{a_A, a_B, a_C, b_A, b_B, b_C, \emptyset\}$ ).

**Table 2**  
Judgment model fits.

	CV	BIC	$\tau_1$	$\tau_2$	Best
<i>Experiment 1</i>					
Random	-5215	10 430			9 (16)
IO	-4151	8259		0.62	28 (26)
IO/N	<b>-4045</b>	<b>8057</b>	0.78	0.32	37 (32)
<i>Experiment 2</i>					
Random	-10 310	20 620			13 (16)
IO	-8030	15 983		0.59	26 (2)
IO/N	<b>-7894</b>	<b>15 644</b>	0.67	0.34	56 (77)

Note: The “best” column displayed the number of individuals best-fit by each model according to CV, with BIC results in the brackets.

While in principle this intervention decision needs to be made constantly, at every instant throughout the trial. In practice we simplified our analyses by assuming that learners make exactly one intervention selection decision per second.<sup>3</sup>

### 5.1. Expected information gain

Information gain (IG) is a common currency for measuring the value of evidence for an ideal learner (Coenen, Nelson, & Gureckis, 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, that best reduces 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.<sup>4</sup> 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 experienced so far:

$$H(S)_{t_x} = \sum_{s \in S} P(s)_{t_x} \log_2 \frac{1}{P(s)_{t_x}} \tag{2}$$

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 atemporal 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  $\bar{\mathbf{o}}$  under different structures. We further assume the number of samples simulated under each structure is based on the structure’s (current) prior probability (Nelson, 2005).<sup>5</sup> For each simulated future  $\bar{\mathbf{o}} \in \bar{\mathbf{o}}$  we compute the information gain as:

$$IG(i^*, \bar{\mathbf{o}})_{t_x}^{t_y} = H(S)_{t_x} - H(S, i^*, \bar{\mathbf{o}})_{t_y} \tag{3}$$

and expected information gain as:

$$EIG[i^*]_{t_x}^{t_y} = \sum_{\bar{\mathbf{o}} \in \bar{\mathbf{o}}} IG(i^*, \bar{\mathbf{o}})_{t_x}^{t_y} \tag{4}$$

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 ( $i^* = \emptyset$ ).

Fig. 14 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 ( $\emptyset$ ) 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 ( $\{b_A, \dots, b_C\}$ ) also have low expected information since, at this stage, they would only serve to block potentially informative dynamics produced by the previous intervention.

<sup>3</sup> 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 within 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.

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

<sup>5</sup> All results are based on sampling 512 event sequences from each window. We selected this number for computational practicability and it is a multiple of 64 which is the cardinality of three-component structures and square root of the cardinality of four-node structures. We checked that this sample size resulted in stable and consistent results by replicating the simulation process with different seeds.

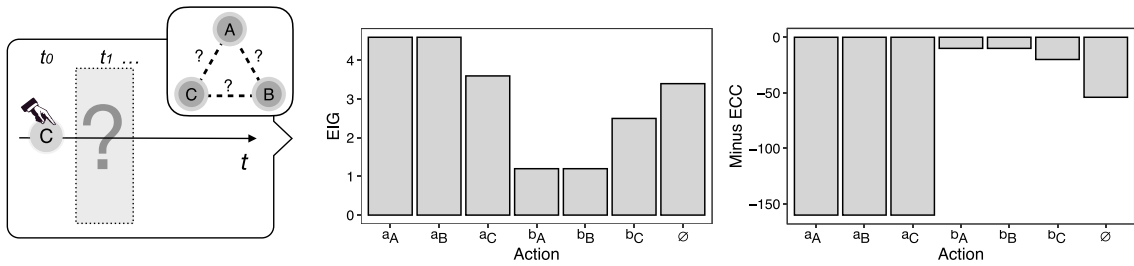


Fig. 14. Example of expected information gain (EIG) and expected computational cost (ECC). The learner activated  $C$  at  $t_0$  and is now deciding what to do at  $t_1$ . The notions of  $a_X$ ,  $b_X$ , and  $\emptyset$  stand for choices to activate  $X$ , block  $X$ , or do nothing, respectively. Both EIG and ECC are temporally discounted. ECC was calculated based on expected local events with a polynomial function.

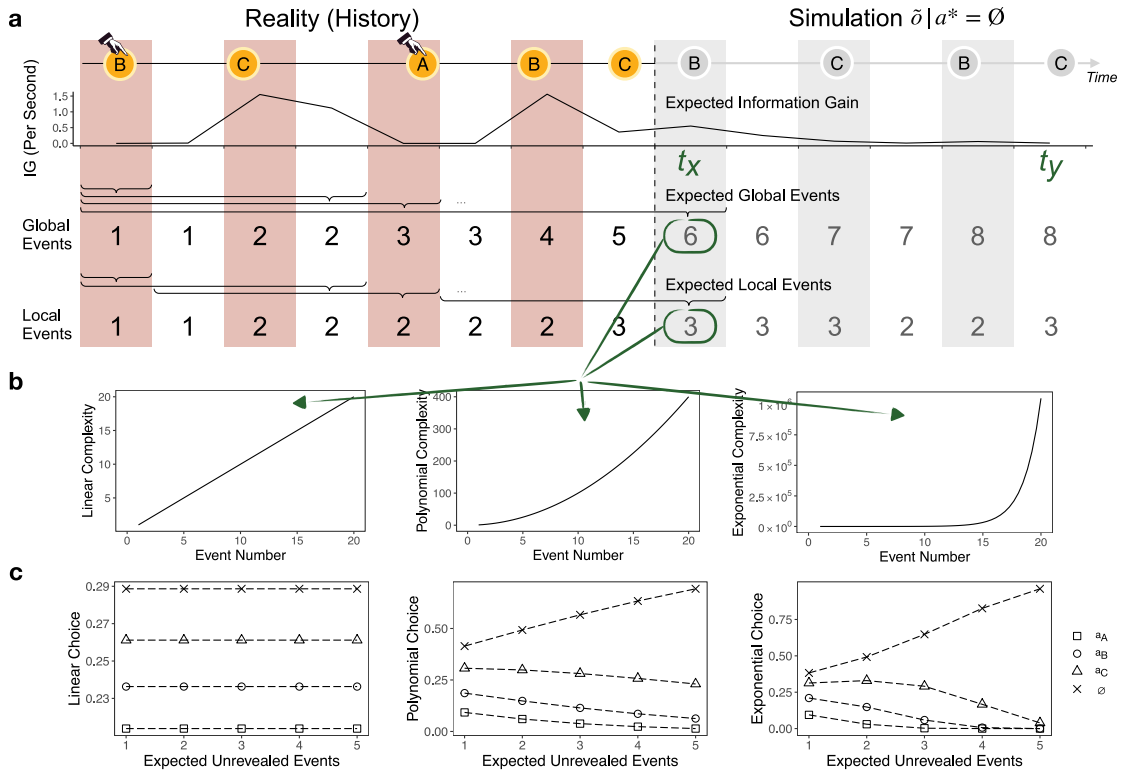


Fig. 15. An illustrative example of historical and *expected* Information Gain (IG), alongside historical and *expected* global events and local events. (a) Pink portion (left of  $t_x$ ) = Window-by-window IG about true structure; Global events (since start of observation); and Local events (within the last 4 s). Gray portion (right of  $t_x$ ) *Expected* upcoming information, global and local events. (b) Three possible computational cost functions of event number. (c) A demo of how different complexity functions react to the number *expected* unrevealed events under the softmax function. *Expected* evidence was generated from a  $A \rightarrow B \rightarrow C$  chain. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.2. *Expected 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 (Fig. 2a), 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” (see Fig. 15a; c.f. Gerstenberg, Bechivanidis, & Lagnado, 2013). That is, we assume learners enumerate and consider causal pathways involving events or interventions from up to several seconds prior to

the moment at which the inference is taking place. We chose 4 s 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 that shifts throughout a trial as a function of the number of recent events (see Fig. 15a). We will also examine other choices of the window size in Table C.1.

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 the same time since this quickly becomes intractable as the number of components increases. For instance, there are 4064 possible structures linking 4 components together and this would increase to 1,048,576 if there were 5 components in the system. 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 succeed in the wild, we think of the idealized Bayesian inference as an upper bound on the computational cost of inference. We 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  $c$ . These functions, including linear  $O(n)$ , polynomial  $O(n^c)$ , exponential  $O(c^n)$  scaling, differ at how fast the cost grows with the increase of event numbers (Fig. 15b).

Similar to expected information gain, we can also anticipate computational cost (CC) of integrating future evidence. This involves counting the events occurring in simulated outcomes  $\bar{o} \in \bar{\mathbf{o}}$ . 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:

$$CC(i^*, \bar{o}_{t_x}^{t_y}) = f_{\text{complexity}}(n_{t_x}^{t_y}, c). \quad (5)$$

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 Fig. 15).

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

$$ECC[i^*]_{t_x}^{t_y} = \sum_{\bar{o} \in \bar{\mathbf{o}}} CC(i^*, \bar{o}_{t_x}^{t_y}) \quad (6)$$

### 5.3. Resource-rational intervention utility

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:

$$\mathbb{E}[U(i^*)] = \sum_{t=t_x}^{t_y-1} R(t) \cdot [\text{EIG}[i^*]_t^{t+1} - \omega \cdot \text{ECC}[i^*]_t^{t+1}] \quad (7)$$

where we assume a 1 s granularity for measurement, and where  $\omega$  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):

$$R(t) = 1 - \frac{t}{t_{\text{end}}} \quad (8)$$

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

$$\operatorname{argmax}_{i^* \in \mathbf{I}} [\mathbb{E}[U(i^*)]] \quad (9)$$

Fig. 15 visualizes the various elements of a trial that combine into our resource rational algorithm and Fig. 14 shows an example in which information gain and inferential complexity differ in the choices they favor and hence trade off. 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.

### 5.4. Model fitting

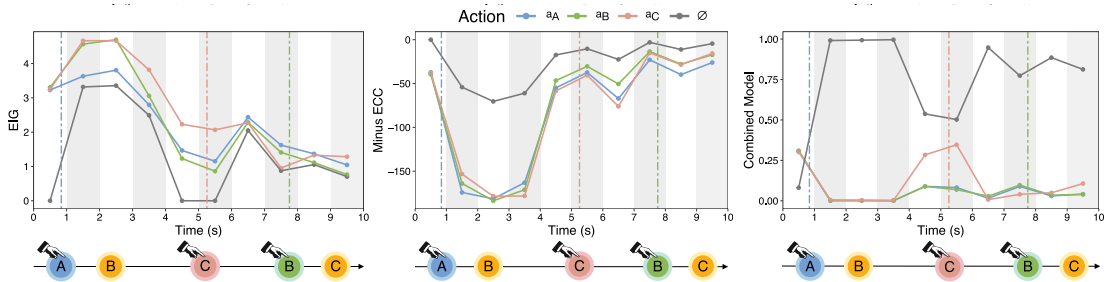
Our primary class of models is based on the utility function specified in Eq. (7). That is, one that is sensitive to both expected information and computational cost. The measurement of this cost is formed as (linear, polynomial, or exponential) complexity

<sup>6</sup> Intuitively, this horizon is reasonable here for several reasons: (1) The rational temporal discount factor makes the 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 complicated causal interactions and compounded by the learner’s uncertainty thereof leads to complicated simulated dynamics whose predictive power rapidly drops toward chance beyond a few seconds (cf. Bramley, Gerstenberg, Tenenbaum, & Gureckis, 2018).

**Table 3**  
Intervention model fits.

	CV	BIC	$\tau$	$\omega$	$\theta$	Best
<i>Experiment 1</i>						
Baseline	-17 033	34 027	0.31			7 (14)
EIG	-16 588	33 109	3.01		10.80	10 (7)
EIG-ECC <sup>Global</sup>	-16 456	32 880	2.67	$4.89 \times 10^{-3}$	8.99	11 (4)
EIG-ECC <sup>Local</sup>	<b>-16 415</b>	<b>32 792</b>	2.27	$8.82 \times 10^{-3}$	7.91	46 (49)
<i>Experiment 2</i>						
Baseline	-43 042	85 997	0.27			2 (3)
EIG	-40 507	80 887	1.98		8.13	19 (51)
EIG-ECC <sup>Global</sup>	-40 507	80 898	1.98	$1.69 \times 10^{-6}$	8.13	14 (2)
EIG-ECC <sup>Local</sup>	<b>-40 462</b>	<b>80 804</b>	1.86	$1.59 \times 10^{-3}$	7.63	60 (39)

Note: The “best” column displayed the number of individuals best-fit by each model according to CV, with BIC results in the brackets. BIC for the fully random baseline was 97 210 in Experiment 1 and 262 515 in Experiment 2. Parameters reported were based on BIC results.



**Fig. 16.** Example of real-time model prediction for a participant in reliable condition of Experiment 1 facing  $A \rightarrow B \rightarrow C$  structure. Lines and points show instantaneous value for each potential intervention (colors) or non-intervention (black). Dashed vertical lines show participants interventions. Model takes earlier interventions and observations into consideration and predicts value of intervention choices for each 1-second window (marked by vertical white/gray shading). Parameters of the combined model based on EIG + local polynomial cost model fit to this individual. Model fit is the product of likelihoods of the chosen action or non-action in each window. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

functions of (global or local) real-time expected event numbers. 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. Note that only polynomial or exponential scaling can capture the phenomenon that the more the expected unrevealed events caused by previous interventions, the greater the likelihood of avoiding future activating interventions (Fig. 15c). We used polynomial scaling with a generic exponent of 2 as the primary form here while the results of other complexity functions with different exponents or bases can be found in Table C.2.

To investigate whether participants are sensitive to both expected information gain and expected computational cost, we also examined a purely information-driven model that removes ECC from Eq. (10). However, a model that considers computational costs could easily beat such a purely information-driven model given the fact that greedy-EIG underestimates the value of waiting when opportunities to intervene are finite.<sup>7</sup> In contrast, the vast majority of time windows in the human data did not contain an action.<sup>8</sup> Therefore, we included a constant bias against action that increased the probability of not acting:  $B(i) = 1$  if  $i = \emptyset$  and  $B(i) = 0$  otherwise. This allows for a fairer comparison between dynamic-cost-dependent and cost-free models. If our EIG-ECC model outperforms the EIG model, this means that participants timed their interventions in a reactive way to cope with the expectation of computational cost rather than simply avoiding action with a constant probability across time.

We assumed stochasticity in participants’ intervention choices captured by a softmax function (Luce, 1959) over the resultant values. The resource-rational prediction is:

$$P(\text{intervention} = i) = \frac{\exp((\mathbb{E}[U(i)] + \theta B)/\tau)}{\sum_{i' \in \mathcal{I}} \exp((\mathbb{E}[U(i')] + \theta B)/\tau)} \tag{10}$$

The model fitting procedure is similar to what we used for the judgment models. We provide hold-one-device-out cross-validation results and BIC results at both the aggregate and individual levels. As shown in Table 3, across both experiments, models that

<sup>7</sup> Conceptually, this is because it does not incorporate the expected utility of saving an action for later use.  
<sup>8</sup> This is also owing to the fact that time spent on activities that were not directly connected to the learning process, such as marking their answers and moving the mouse, was taken into account.

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 (78% of participants were fit best by one of the cost-dependent models, 63% people were fit by the local cost model specifically). Fig. 16 gives an example from Experiment 1 in which the combination of expected information and cost give a better account of participants' intervention choices than either does alone. In Appendix B, we used the fitted parameters to simulate interventions and judgments and showed that these align qualitatively with the human patterns.

The fact that the boundedly rational models outperformed pure information seeking models corroborates our central idea that participants' interventions were shaped both by how much information they expected to gain and by how hard they would have to work to process potential future information. Note that while we did not present them for space reasons, model variants sensitive to only cost but not information, perform worse than all the models we present irrespective of how the cost is calculated. These models invariably favor waiting or blocking over activating components.

More individuals in Experiment 2 were best 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 be due to the fact 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 this greedy model.

### 5.5. Prospective vs. retrospective complexity

We explored whether expected computational cost — which depended on both recent events but also how many events are anticipated to happen in the near future — can be substituted with a simpler retrospective computational cost consideration — based only on how many events have occurred recently. To test this, we ran retrospective variants of each model in Table C.3 finding that these were always a slightly worse fit than their prospective versions. This could be because, while the retrospective approach captures a sensible and simple heuristic of waiting until the system is quiet, this behavior can also be accounted for by expected complexity. Moreover, retrospective complexity is insensitive to earlier learning about the structure within a trial. For instance, one might have learned that the current system is highly excitatory (captured by the evolving prior shaping expected complexity) but that activity might have died down by the time of the next intervention.

Retrospective complexity also only accounts for when to intervene but not where. In contrast, expected complexity serves as a guide for both when and where to intervene. For example, if a learner has already discovered a component that seems to generate a large number of events, they may decide not to activate it again. To test whether cost-dependent choice of where to intervene is a significant feature of our participants' intervention selections, we also fit resource-rational models to only the time windows where participants made activations. At the aggregate level, the EIG-ECC model (cross-validated log-likelihood:  $-4788$  in Experiment 1 and  $-9054$  in Experiment 2) only had minor advantages over a pure EIG model in predicting these windows (cross-validated log-likelihood:  $-4789$  in Experiment 1 and  $-9054$  in Experiment 2). However, it did also better capture 31% of individuals. The individuals best-fit by the EIG-ECC model in terms of the windows in which they acted had better performance than 55% people who were best-fit by an EIG only model (accuracy: 67% vs. 59%,  $t(144) = 2.22$ ,  $p = .028$ , Cohen's  $d = 0.38$ ), and better performance than the 14% people best-fit as selecting components to intervene on randomly (accuracy: 67% vs. 49%,  $t(74) = 3.97$ ,  $p < .001$ , Cohen's  $d = 0.99$ ).

## 6. General discussion

In a dynamically unfolding world, uncovering causal relationships requires online control and processing of continuous sensory information. To learn about how the world works, one must choose *where* to act, *how* to act, and *when* to do so while also tracking what happens before, during, and after one's actions. In this paper, we investigated human learning in a setting where learners use freely timed interventions to investigate the underlying causal structure responsible for devices' patterns of real-time component activations. We investigated what factors affected the quality of their inferences, and what strategies they used to choose and time their interventions. We hypothesized that computational limitations, and a rational anticipation thereof, would play a key 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 and used model fitting to show that this could help explain both human judgments and intervention patterns.

### 6.1. What we found

Our empirical findings fall into two classes: (1) Insights about features of real time causal systems that determine how easily people can learn them and (2) insights about how people choose interventions to support their learning.



*What experimental factors affected participants' learning?* Across both our experiments, participants had more success identifying acyclic than cyclic structures while an ideal observer model showed the reverse pattern, highlighting a fundamental divergence between ideal and bounded learning. The ideal observer benefits from the higher density of events produced by feedback loops, essentially because it is able to enumerate and marginalize over the many possible causal explanations for the data, and make ideal use of the rich timing information. Meanwhile, participants' ability to do this was presumably limited by their information processing capacity, leading to a kind of "less is more" phenomenon (cf. Gigerenzer & Todd, 1999) in which simpler evidence was often more valuable to them even when less normatively informative. In line with this, we showed that human accuracy patterns can be accounted for through a corruption of the ideal observer that assumed bandwidth limitations on the processing of evidence, such that inferential noise and probability of error increase with the compounding effect of event density, potentially more than counteracting the value of the additional information.

While past work has demonstrated that people are sensitive to delay reliability, and use delay information in addition to order to shape causal structure judgments (Bramley, Gerstenberg, Mayrhofer, & Lagnado, 2018; Greville & Buehner, 2010), reliability had little impact on performance here. Delay condition did not make a statistically significant difference to accuracy in either experiment suggesting reliable delays may be less critical in the active learning setting, where interventions can be dynamically adjusted to accommodate experienced variability. Another possibility is that our delay manipulation was too subtle. However, our unreliable condition had a sevenfold wider standard deviation which is both salient in visualizations of the trials and commensurate with past work (Bramley, Mayrhofer, et al., 2017). We note also that the top 10% of performers were almost all in the reliable condition (100% in Experiment 1 and 90% in Experiment 2). We take this to suggest that delay reliability is important for achieving high accuracy. Additionally, reliable condition participants were better at identifying that full-loop cyclic structures than unreliable condition participants which might suggest that reliable delays allowed extended temporal patterns like periodicity of cyclic activations to contribute to structure identification.

We found several other systematic judgment errors. Some participants mistook chain structures as fully-connected, marking extraneous indirect links from the root components to distal child components. This lines up with the findings of a number of atemporal causal learning studies (Fernbach & Sloman, 2009; Lagnado & Sloman, 2004) as well as studies that have used continuous valued variables (Davis et al., 2020). This pattern may reflect the "local computations" idea that people often focus at subparts of the larger system, such as on pairs of variables, and so experience an appearance of direct causation when observing indirectly connected components. Participants were also quite likely to mistake fork structures as chain structures. Since the outputs of a fork would invariably occur in some staggered pattern, this seems straightforwardly consistent with occasional fallback on a heuristic of taking temporal order to directly reflect causal order (cf. Bramley, Gerstenberg, Mayrhofer, & Lagnado, 2018; McCormack et al., 2016).

Among the cyclic structures, participants frequently mistook output components as constituent elements of their upstream feedback loops (e.g. in *Cyclic3*, *Cyclic5* and *Cyclic6*). This is not unreasonable because these output components tended to activate almost in lockstep with components of the feedback loop that produced them. While informative to an ideal observer, the subtle differences in inter-event delays between components that formed part of the loop, and components that formed the loop's outputs were presumably difficult to process for bounded human learners.

*How do people choose what interventions to perform?* Participants generally performed fewer activations in cyclic structures. Multiple activating interventions in cyclic structures could quickly compound the complexity of the subsequent evidence, presumably overwhelming bounded learners and motivating learners to avoid this. For four-node structures, more activations were required to achieve an equivalent level of certainty as in three-node structures, and participants performed more activations on these problems especially in Experiment 1. More generally, participants appeared to adjust their interventions depending on the current and anticipated event density. They tended to perform more activating interventions at moments when the number of expected unrevealed events was low, and were more likely to wait or use blocks to reset the system when the number of expected unrevealed events was high. Strikingly, in both experiments, the participants who performed well were those who managed to generate evidence that was relatively more informative and less complex (Fig. 7).

In Experiment 2, participants did not use blocks nearly as often as activations, but most participants did use them and did so particularly in excitable structures and at moments of high anticipated complexity. Participants performed better on the full-loop structures when allowed to block, compared to Experiment 1 (*Cyclic1* and *Cyclic4*). Full-loop structures are intuitively far easier to understand when exhibiting a single oscillating cycle of activity. If participants performed a second activation intervention before such a cycle died out, they would face confusing evidence patterns with two excitations traveling around the system in tandem, potentially overtaking one another and going in and out of sync. Resetting blocks in particular seemed to help learners control such complicated scenarios.

## 6.2. 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, Goodman, Griffiths, & Tenenbaum, 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 (Callaway et al., 2022), information sampling (Petitet, Attaallah, Manohar, & Husain, 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. By building and comparing computational models, we firstly showed that participants' causal judgments depended on both the informativeness and the complexity of the generated evidence. More importantly, we then 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 processable to be useful.

Specifically, out of the measures of complexity we examined, a polynomial function of inference-relevant events (those in the recent past and expected in the near future) best captured the influence of complexity on intervention choice, and we found that prospective complexity as well as retrospective complexity contributed to participants' choices. 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 inferences, we feel this reflects an intuitively sensible and plausible sensitivity to local events capturing the fact that inferential complexity compounds as the number of actual causal relata increase (Bramley, Dayan, et al., 2017; Fernbach & Sloman, 2009; Van Rooij, Blokpoel, Kwisthout, & Wareham, 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), these are typically intractable, requiring a run time that scales often far worse than exponentially 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, 2022), 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 incrementalist accounts. It has been argued that learners form one or a few hypotheses at a time (Bonawitz et al., 2014), or focus 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 the atemporal 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 total number of components. Indeed, the run time of 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.

### 6.3. Insights for a process-level model of real time active causal learning

*Causal structure induction.* We have shown how complexity effects human judgments from a computational level perspective, essentially shaping the nature of the optimization problem faced by active causal learners in real time (Marr, 1982). Recent work on observational causal learning has generally found simple event *order* to be a strong driver of structure judgments, with delay expectations tending to have smaller and subtler effects on inferences (Bramley, Gerstenberg, Mayrhofer, & Lagnado, 2018; Valentin et al., 2020). As such, some form of a simpler endorser heuristic (Bramley et al., 2015; Fernbach & Sloman, 2009) which assumes that people attribute an effect to the most recent event may do a reasonable job of describing some participants judgments. Such a heuristic, however, does not explain participants' good performance in identifying acyclic and cyclic classes. Taking fork structures for example (Acyclic3 and Acyclic6), the order of child node activations would often reverse following repeated activations of the root node, but participants rarely reported bidirectional connections between these child components. More fundamentally, a learner that relies on a temporal order heuristic would lack the necessary representation of the hypothesis space and uncertainty needed to guide intervention selection. We would need a separate account for how people make intervention choices, and a linking model for how the inferences and activation choices are connected with one another.

One other finding that a comprehensive process account would need to accommodate, is that in cyclic structures, participants frequently drew a bidirectional link between the output component and the component in the loop that tended to activate closest in time (Cyclic3, 5--7, and 8), despite the fact that participants were trained to expect longer delays than this between truly causally related events. One interpretation is that repeated spontaneous activations are a prototypical signal of cyclic relationships (Valentin et al., 2020). Our participants may have reasoned in terms of this abstract feature, and applied it without thinking more carefully about the exact timing at which the events happened (Goodman, Ullman, & Tenenbaum, 2011). Alternatively it could be that our pre-training on delays was not sufficient to overrule a prior expectation of causal contiguity, that is, that cause and effect events happen close in time (Hume, 1740). Indeed, past studies have required strong manipulations to override this preference (Buehner & May, 2004; Buehner & McGregor, 2006). More work is needed to better understand how temporal delays affect causal judgments.

**Intervention.** One long-standing debate centers on the question of whether human active learning and 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 might repeat-test it, or probe the components at which the effects occurred. In this way, learners might follow a kind of extended positive testing strategy in which they focus their energies on components deemed to be to produce effects so as to gather evidence “by making the machine go”. This reflects the rationale behind positive testing 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, distinct from 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.

Additionally, 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 computational norms. However, the current work shows that whatever heuristic or adaptive toolbox is proposed, to fully capture intervention choice, it must include strategies that are at least be somewhat responsive to experienced and anticipated complexity.

#### 6.4. Future questions

Some questions related to continuous-time causal learning remain for future research. One open question is where *cyclic structures* fit into the overall landscape of causal cognition. A full representation of a cyclic system seems to demand a temporal dimension and predictions are generally sensitive to the system’s current state. In our experiments, participants performed well above chance for most connections in most cyclic structures without extra training, and were able to 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. Nevertheless, some cyclic structures may be particularly challenging for people to understand, and the reasons for this may go beyond what is captured by our general computation cost account. For example, participants performed relatively poorly in identifying the internal structure of cyclic structures with an output of a loop as mentioned in the paragraph above. 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. This makes it valuable to explore more factors that affect learning and control in cyclic systems.

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 discretization. Furthermore, interventions might help to create data that is more amenable to these forms of summarization. Better understanding of human causal induction requires us to go beyond covariation-based causal Bayesian networks (Pearl, 2000), but this should not involve discarding 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.

#### 6.5. Conclusions

Everyday experience is rich with events that reoccur and can be causally related in ways that allow us to predict, control, and make sense of what has happened and what is likely to happen next. While previous research on active causal learning has often 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 gathering and interpreting that evidence. Our experiments and modeling show that participants’ causal judgments depend on not just the informativeness but also the complexity of evidence they gather, and that they adapt their actions to the ongoing event dynamics during learning so as to strike a balance between expected information gain and anticipated inferential complexity. 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 research into human learning.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

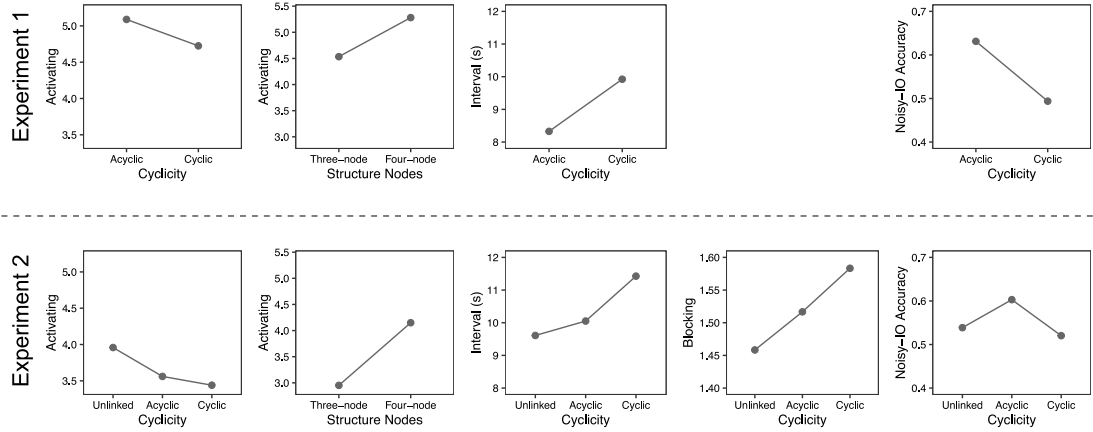


Fig. B1. Results from simulated evidence according to the parameters fit in the intervention and judgment models.

## Data availability

The experimental procedure, data, and analysis code are available at: [https://github.com/tianweigong/time\\_and\\_intervention](https://github.com/tianweigong/time_and_intervention).

## Appendix A. Ideal observational (io) learning

We formulate how a learner should ideally update their beliefs after seeing evidence produced by 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:

$$P(S|\mathbf{d};\mathbf{i}) \propto p(\mathbf{d}|S;\mathbf{i}) \cdot P(S). \quad (11)$$

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  $d_{\text{component}}^{(\text{index})}$  and that this is conditioned on the set of interventions  $\mathbf{i}$  including all activations  $a_{\text{component}}^{(\text{index})}$  and blocks  $b_{\text{component}}^{(\text{index})}$  performed by the learner during the learning episode.

As mentioned in the main text, when calculating the likelihood of the data given a candidate structure, there are likely to be multiple potential paths of actual causation. Each of these has its own likelihood. To construct the total likelihood of a hypothesized causal structure and interventions producing a set of events, we must consider all possible causal paths  $Z_s$  that could describe what actually happened given structure  $s$  and then repeat this for every  $s \in S$ . Since the path set 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:

$$p(\mathbf{d}|s;\mathbf{i}) = \sum_{z' \in Z_s} p(\mathbf{d}|z';\mathbf{i}). \quad (12)$$

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

$$\gamma_{\text{pdf}}(t_{\text{effect}} - t_{\text{cause}}, \alpha, \beta). \quad (13)$$

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<sup>9</sup>:

$$w \underbrace{\left[ \gamma_{\text{cdf}}(\max(0, t_{\text{block\_onset}} - t_{\text{cause}}), \min(t_{\text{block\_offset}} - t_{\text{cause}}, t_{\text{now}} - t_{\text{cause}}), \alpha, \beta) \right]}_{\text{activation was blocked}} + \quad (14)$$

$$\underbrace{\left[ (1 - \gamma_{\text{cdf}}(t_{\text{now}} - t_{\text{cause}}, \alpha, \beta)) \right]}_{\text{activation has not occurred yet}} + \underbrace{(1 - w)}_{\text{activation failed}} \quad (15)$$

Thus, the likelihood for a particular causal path given a particular causal structure can be calculated exactly via a combination of

<sup>9</sup>  $\gamma_{\text{cdf}}(x, y, \alpha, \beta)$  in Eq. (15) denotes the cumulative probability of a delay being between  $x$  and  $y$  in length.

**Table C.1**  
Intervention model fits of back tracking window size in polynomial local cost models with a generic exponent of 2.

Window	CV	BIC	$\tau$	$\omega$	$\theta$
<i>Experiment 1</i>					
1 s	-16 426	32 810	2.21	$5.38 \times 10^{-2}$	7.77
2 s	-16 430	32 797	2.23	$1.94 \times 10^{-2}$	7.82
3 s	-16 414	32 789	2.24	$1.17 \times 10^{-2}$	7.82
5 s	-16 418	32 798	2.30	$7.30 \times 10^{-3}$	7.99
6 s	-16 421	32 805	2.34	$6.46 \times 10^{-3}$	8.09
7 s	-16 425	32 814	2.38	$6.12 \times 10^{-3}$	8.21
<i>Experiment 2</i>					
1 s	-40 466	80 811	1.86	$8.81 \times 10^{-3}$	7.66
2 s	-40 459	80 798	1.86	$3.42 \times 10^{-3}$	7.63
3 s	-40 457	80 794	1.85	$2.17 \times 10^{-3}$	7.60
5 s	-40 467	80 813	1.86	$1.29 \times 10^{-3}$	7.65
6 s	-40 471	80 821	1.87	$1.11 \times 10^{-3}$	7.67
7 s	-40 477	80 833	1.88	$9.74 \times 10^{-4}$	7.72

**Table C.2**  
Intervention model fits of exponents or base parameters in polynomial- or exponential-costs.

Experiment 1						
	CV	BIC	$c_{CV/BIC}$	$\tau$	$\omega$	$\theta$
<i>Global cost</i>						
EIG-ECC <sub>Linear</sub>	-16 550	33 040		2.00	$1.00 \times 10^{-1}$	6.80
EIG-ECC <sub>Polynomial</sub>	-16 451	32 882	1.8	2.55	$1.03 \times 10^{-2}$	8.52
EIG-ECC <sub>Exponential</sub>	-16 550	33 067	2.4/1.8	2.79	$1.44 \times 10^{-10}$	10.00
<i>Local cost</i>						
EIG-ECC <sub>Linear</sub>	-16 524	32 994		1.82	$1.32 \times 10^{-1}$	6.13
EIG-ECC <sub>Polynomial</sub>	-16 378	32 743	1.6/1.4	1.80	$6.22 \times 10^{-2}$	6.04
EIG-ECC <sub>Exponential</sub>	-16 556	33 083	1.8	2.79	$1.39 \times 10^{-10}$	10.02
Experiment 2						
	CV	BIC	$c_{CV/BIC}$	$\tau$	$\omega$	$\theta$
<i>Global cost</i>						
EIG-ECC <sub>Linear</sub>	-40 507	80 898		1.98	$3.33 \times 10^{-6}$	8.13
EIG-ECC <sub>Polynomial</sub>	-40 504	80 901	3	1.96	$4.40 \times 10^{-6}$	8.06
EIG-ECC <sub>Exponential</sub>	-40 503	80 895	1.2/1.18	1.97	$7.31 \times 10^{-11}$	8.09
<i>Local cost</i>						
EIG-ECC <sub>Linear</sub>	-40 507	80 898		1.98	$3.76 \times 10^{-6}$	8.13
EIG-ECC <sub>Polynomial</sub>	-40 462	80 815	2	1.86	$1.59 \times 10^{-3}$	7.63
EIG-ECC <sub>Exponential</sub>	-40 500	80 896	1.18/3.2	1.97	$8.65 \times 10^{-11}$	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 given a fixed  $c$ . We reported two  $c$  with CV and BIC deviated in their results.

**Table C.3**  
Retrospective local complexity model fits.

	CV	BIC	$\tau$	$\omega$	$\theta$
<i>Experiment 1</i>					
EIG-ECC <sub>Linear</sub>	-16 566	33 077	3.11	$3.30 \times 10^{-1}$	10.83
EIG-ECC <sub>Polynomial</sub>	-16 588	32 120	3.01	$1.67 \times 10^{-5}$	10.80
EIG-ECC <sub>Exponential</sub>	-16 620	33 120	3.00	$1.57 \times 10^{-7}$	10.80
<i>Experiment 2</i>					
EIG-ECC <sub>Linear</sub>	-40 507	80 898	1.98	$4.61 \times 10^{-5}$	8.13
EIG-ECC <sub>Polynomial</sub>	-40 506	80 898	1.98	$4.74 \times 10^{-6}$	8.12
EIG-ECC <sub>Exponential</sub>	-40 507	80 898	1.98	$3.01 \times 10^{-11}$	8.13

Note: To fit retrospective models, we replaced the complexity component in Eq. (7) with a retrospective recent event count. This assigned a cost according to the number of events in previous 4 s for activation interventions  $\{a_A, a_B, a_C, a_D\}$ , while assigned a cost of zero for doing nothing or blocking interventions.

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.

## Appendix B. Comparing simulated resource-rational interventions and judgments

To test whether our intervention and judgment models can replicate participants' qualitative behavior patterns, we use the parameters fit by human data to simulate resource-rational agents that intervene on the same devices examined in Experiments 1 and 2. In both the reliable and unreliable delay conditions, we generated 30 simulated learners. The intervention patterns are shown in Fig. B1. This shows that for both experiments, simulated resource-rational learners activated components in acyclic structures more than cyclic structures, and four-node structures more than three-node structures. They waited longer to perform their next intervention when the structures were cyclic then when they were acyclic. For Experiment 2, they performed more blocking actions in cyclic devices than acyclic devices. These results demonstrate that our intervention model is capable of replicating a wide range of human intervention patterns.

We also provided simulated evidence to the judgment model, which was based on parameters fitted with human data. For both experiments, the noisy-IO judgment model replicated the human result that acyclic structures had higher accuracy than cyclic structures. Unlike participants, these simulations were not more accurate on unlinked structures in Experiment 2. This could be due to some extra assumptions that we did not include in our models, such as the possibility that rather than beginning each trial with a uniform prior over structures, participants may have expected causal models to be sparse (Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2008).

## References

- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Psychology Press, <http://dx.doi.org/10.4324/9780203771730>.
- Austerweil, J. L., & Griffiths, T. L. (2011). Seeking confirmation is rational for deterministic hypotheses. *Cognitive Science*, 35(3), 499–526. <http://dx.doi.org/10.1111/j.1551-6709.2010.01161.x>.
- 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. <http://dx.doi.org/10.1016/j.cogpsych.2014.06.003>.
- 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. <http://dx.doi.org/10.1037/rev0000061>.
- 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. <http://dx.doi.org/10.1037/xlm0000548>.
- Bramley, N. R., Gerstenberg, T., Tenenbaum, J. B., & Gureckis, T. M. (2018). Intuitive experimentation in the physical world. *Cognitive Psychology*, 195, 9–38. <http://dx.doi.org/10.1016/j.cogpsych.2018.05.001>.
- Bramley, N. R., Jones, A., Gureckis, T. M., & Ruggeri, A. (2022). Children's failure to control variables may reflect adaptive decision-making. *Psychonomic Bulletin & Review*, <http://dx.doi.org/10.3758/s13423-022-02120-1>.
- 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. <http://dx.doi.org/10.1037/xlm0000061>.
- 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).
- Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, 23(3), 389–411. <http://dx.doi.org/10.1037/met0000159>.
- 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. <http://dx.doi.org/10.1037/0278-7393.29.6.1119>.
- Buehner, M. J., & May, J. (2004). Abolishing the effect of reinforcement delay on human causal learning. *Quarterly Journal of Experimental Psychology Section B*, 57(2), 179–191. <http://dx.doi.org/10.1080/02724990344000123>.
- Buehner, M. J., & McGregor, S. (2006). Temporal delays can facilitate causal attribution: Towards a general timeframe bias in causal induction. *Thinking & Reasoning*, 12(4), 353–378. <http://dx.doi.org/10.1080/13546780500368965>.
- Burns, P., & McCormack, T. (2009). Temporal information and children's and adults' causal inferences. *Thinking & Reasoning*, 15(2), 167–196. <http://dx.doi.org/10.1080/13546780902743609>.
- Callaway, F., van Opheusden, B., Gul, S., Das, P., Krueger, P. M., Lieder, F., et al. (2022). Rational use of cognitive resources in human planning. *Nature Human Behaviour*, <http://dx.doi.org/10.1038/s41562-022-01332-8>.
- 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. <http://dx.doi.org/10.1111/1467-8624.00081>.
- Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, 39, 1–72. <http://dx.doi.org/10.1017/S0140525X1500031X>.
- 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. <http://dx.doi.org/10.3758/s13423-018-1470-5>.
- Coenen, A., Rehder, B., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology*, 79, 102–133. <http://dx.doi.org/10.1016/j.cogpsych.2015.02.004>.
- 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. <http://dx.doi.org/10.1037/xlm0000680>.
- Dasgupta, I., Schulz, E., & Gershman, S. J. (2017). Where do hypotheses come from? *Cognitive Psychology*, 96, 1–25. <http://dx.doi.org/10.1016/j.cogpsych.2017.05.001>.
- Davis, Z., Bramley, N. R., & Rehder, B. (2020). Causal structure learning in continuous systems. *Frontiers in Psychology*, 11, 244. <http://dx.doi.org/10.3389/fpsyg.2020.00244>.
- 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. <http://dx.doi.org/10.1016/j.cogpsych.2018.01.002>.
- Fernbach, P. M., & Sloman, S. A. (2009). Causal learning with local computations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(3), 678–693. <http://dx.doi.org/10.1037/a0014928>.
- Gershman, S. J. (2020). Origin of perseveration in the trade-off between reward and complexity. *Cognition*, 204, Article 104394. <http://dx.doi.org/10.1016/j.cognition.2020.104394>.

- Gerstenberg, T., Bechlivanidis, C., & Lagnado, D. A. (2013). Back on track: Backtracking in counterfactual reasoning. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual conference of the cognitive science society*. (pp. 2386–2391).
- 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. <http://dx.doi.org/10.1037/rev0000281>.
- Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press.
- 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 42nd annual conference of the cognitive science society* (pp. 2908–2914).
- Gong, T., & Bramley, N. R. (2022). Continuous time causal structure induction with prevention and generation. PsyArXiv. <http://dx.doi.org/10.31234/osf.io/z586q>.
- Goodman, N. D., Ullman, T. D., & Tenenbaum, J. B. (2011). Learning a theory of causality. *Psychological Review*, 118(1), 110–119. <http://dx.doi.org/10.1037/a0021336>.
- 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. <http://dx.doi.org/10.1037/0012-1649.37.5.620>.
- Greville, W. J., & Buehner, M. J. (2010). Temporal predictability facilitates causal learning. *Journal of Experimental Psychology: General*, 139(4), 756–771. <http://dx.doi.org/10.1037/a0020976>.
- 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. <http://dx.doi.org/10.1111/tops.12142>.
- Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review*, 116(4), 661–716. <http://dx.doi.org/10.1037/a0017201>.
- 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. <http://dx.doi.org/10.3758/BF03194330>.
- 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), Article e12926. <http://dx.doi.org/10.1111/cogs.12926>.
- Hume, D. (1740). *A treatise of human nature*. New York: Oxford University Press (2000 reprint).
- 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. <http://dx.doi.org/10.1037/0096-3445.131.4.451>.
- 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. <http://dx.doi.org/10.1037/0278-7393.15.4.596>.
- Kuhn, D., & Brannock, J. (1977). Development of the isolation of variables scheme in experimental and "natural experiment" contexts. *Developmental Psychology*, 13(1), 9–14. <http://dx.doi.org/10.1037/0012-1649.13.1.9>.
- 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. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(4), 856–876. <http://dx.doi.org/10.1037/0278-7393.30.4.856>.
- 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. <http://dx.doi.org/10.1037/0278-7393.32.3.451>.
- 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). New York: Oxford University Press.
- Lai, L., & Gershman, S. J. (2021). Policy compression: An information bottleneck in action selection. In *The psychology of learning and motivation* (pp. 195–227). Academic Press, <http://dx.doi.org/10.1016/bs.plm.2021.02.004>.
- 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. <http://dx.doi.org/10.1017/S0140525X1900061X>.
- 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. <http://dx.doi.org/10.3758/s13423-017-1286-8>.
- Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review*, 115(4), 955–984. <http://dx.doi.org/10.1037/a0013256>.
- 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. <http://dx.doi.org/10.1016/j.cognition.2013.12.010>.
- Lucas, C. G., & Griffiths, T. L. (2010). Learning the form of causal relationships using hierarchical Bayesian models. *Cognitive Science*, 34(1), 113–147. <http://dx.doi.org/10.1111/j.1551-6709.2009.01058.x>.
- Luce, R. D. (1959). *Individual choice behavior*. Hoboken: Wiley.
- Malthus, T. R. (1872). *An essay on the principle of population*. Reeves & Turner.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge: MIT Press, <http://dx.doi.org/10.7551/mitpress/9780262514620.001.0001>.
- McCormack, T., Bramley, N. R., Frosch, C., Patrick, F., & Lagnado, D. (2016). Children's use of interventions to learn causal structure. *Journal of Experimental Child Psychology*, 141, 1–22. <http://dx.doi.org/10.1016/j.jecp.2015.06.017>.
- 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. R., & 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* (pp. 762–767).
- Nelson, J. D. (2005). Finding useful questions: on Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, 112(4), 979–1000. <http://dx.doi.org/10.1037/0033-295X.112.4.979>.
- 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. <http://dx.doi.org/10.1111/bjop.12113>.
- 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*. New York: Cambridge University Press (2009 reprint), <http://dx.doi.org/10.1017/CBO9780511803161>.
- Petitot, P., Attaallah, B., Manohar, S. G., & Husain, M. (2021). The computational cost of active information sampling before decision-making under uncertainty. *Nature Human Behaviour*, 5, 935–946. <http://dx.doi.org/10.1038/s41562-021-01116-6>.
- Rehder, B. (2014). Independence and dependence in human causal reasoning. *Cognitive Psychology*, 72, 54–107. <http://dx.doi.org/10.1016/j.cogpsych.2014.02.002>.
- Rehder, B. (2017). Reasoning with causal cycles. *Cognitive Science*, 41, 944–1002. <http://dx.doi.org/10.1111/cogs.12447>.
- Rottman, B. M., & Hastie, R. (2014). Reasoning about causal relationships: Inferences on causal networks. *Psychological Bulletin*, 140(1), 109–139. <http://dx.doi.org/10.1037/a0031903>.

- 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. <http://dx.doi.org/10.1016/j.cogpsych.2016.05.002>.
- Rottman, B. M., & Keil, F. C. (2012). Causal structure learning over time: Observations and interventions. *Cognitive Psychology*, *64*(1–2), 93–125. <http://dx.doi.org/10.1016/j.cogpsych.2011.10.003>.
- Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science*, *10*(3), 322–332. <http://dx.doi.org/10.1111/j.1467-7687.2007.00587.x>.
- 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. <http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., et al. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*, 99–124. <http://dx.doi.org/10.1146/annurev-neuro-072116-031526>.
- Simon, H. A. (1982). *Models of bounded rationality: empirically grounded economic reason*. MIT Press, <http://dx.doi.org/10.7551/mitpress/4711.001.0001>.
- Slooman, S. A., Love, B. C., & Ahn, W.-K. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, *22*(2), 189–228. [http://dx.doi.org/10.1016/S0364-0213\(99\)80039-1](http://dx.doi.org/10.1016/S0364-0213(99)80039-1).
- Sobel, D. M., & Kushnir, T. (2006). The importance of decision making in causal learning from interventions. *Memory & Cognition*, *34*(2), 411–419. <http://dx.doi.org/10.3758/BF03193418>.
- Soo, K. W., & Rottman, B. M. (2018). Causal strength induction from time series data. *Journal of Experimental Psychology: General*, *147*(4), 485–513. <http://dx.doi.org/10.1037/xge0000423>.
- Spirites, P., Glymour, C. N., Scheines, R., & Heckerman, D. (2000). *Causation, prediction, and search* (2nd ed.). MIT Press, <http://dx.doi.org/10.1007/978-1-4612-2748-9>.
- Stephan, S., Mayrhofer, R., & Waldmann, M. R. (2020). Time and singular causation—A computational model. *Cognitive Science*, *44*(7), Article e12871. <http://dx.doi.org/10.1111/cogs.12871>.
- 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. <http://dx.doi.org/10.1037/xge0001016>.
- Steyvers, M., Tenenbaum, J. B., Wagenmakers, E.-J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*(3), 453–489. [http://dx.doi.org/10.1207/s15516709cog2703\\_6](http://dx.doi.org/10.1207/s15516709cog2703_6).
- 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), 410–441. <http://dx.doi.org/10.1037/rev0000052>.
- Tong, S., & Koller, D. (2001). Active learning for structure in Bayesian networks. In *International joint conference on artificial intelligence*. Vol. 17 (1), (pp. 863–869). Citeseer.
- Ullman, T. D., Stuhlmüller, A., Goodman, N. D., & Tenenbaum, J. B. (2018). Learning physical parameters from dynamic scenes. *Cognitive Psychology*, *104*, 57–82. <http://dx.doi.org/10.1016/j.cogpsych.2017.05.006>.
- 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 42nd 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, <http://dx.doi.org/10.1017/9781107358331>.
- 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. <http://dx.doi.org/10.1111/cogs.12101>.
- White, P. A. (2006). How well is causal structure inferred from cooccurrence information? *European Journal of Cognitive Psychology*, *18*(3), 454–480. <http://dx.doi.org/10.1080/09541440500264861>.
- Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, *27*(2), 172–223. <http://dx.doi.org/10.1016/j.dr.2006.12.001>.