Swipe and hold

Citation for published version:
Btesh, V, Bramley, NR, Speekenbrink, M & Lagnado, DA 2023, Swipe and hold: Composing interventions in continuous time causal learning. in M Goldwater, F Anggoro, B Hayes & D Ong (eds), *Proceedings of the 45th Annual Meeting of the Cognitive Science Society.* vol. 45, Proceedings of the Cognitive Science Society, The Cognitive Science Society, pp. 2370-2376, 45th Annual Meeting of the Cognitive Science Society, Sydney, New South Wales, Australia, 26/07/23. <https://escholarship.org/uc/item/3sv1d2c6>

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Proceedings of the 45th Annual Meeting of the Cognitive Science Society

Swipe and hold: composing interventions in continuous time causal learning

Victor J. Btesh[1*], Neil R. Bramley[2], Maarten Speekenbrink[1], David A. Lagnado[1]
[1]University College London, [2]The University of Edinburgh *victor.btesh.19@ucl.ac.uk

Abstract

We explore intervention policies chosen by participants when learning about causal structures in continuous systems (a demo can be found here[1]). We used the continuous time and state space formalism of Ornstein-Uhlenbeck networks (Davis et al., 2020b). We find that participants' interventions can be understood as compositions of two basic actions: *swipes*, large movements generating large changes in values in a short time window, and *holds*, fixing variables for a few frames and observing the outcome. We further show that intervention policies depend on the underlying causal structure that participants are attempting to retrieve. Swipes were performed more often when there was no outgoing link, suggesting that rapid successive movements was useful in validating the absence of causal relations. On the other hand, interventions featuring longer holding times were favoured in cases where there was at least one outgoing link from the intervened node and appeared to serve the purpose of establishing the sign and strength of causal relationships. Overall, these findings indicate that causal inference in continuous contexts appears to be a two step process, where participants use different interventions for causal discovery than they do to qualify the nature of discovered links, hinting at a tendency to decompose inference to simplify it. [2]

Keywords: active learning; causal reasoning; interventions; continuous time

Introduction

Human learning is a process of information collection and integration. Our bodies allow us not only to be passive learners, being fed data and updating beliefs like most machine learning algorithms, but also active learners, moving or proactively interacting with the world to collect bespoke information. We can choose to gaze in a direction should we feel uncertain about a far away object on the floor. We can physically move towards it, pick up the object, manipulate it, play with it. This active form of learning is ubiquitous in all of us. Children from a young age conduct causally driven experiments to understand the world's dynamics (McCormack et al., 2016; Gopnik et al., 2004; Sobel et al., 2004). People are very efficient observational learners (Rothe et al., 2018) but excellent active learners, dexterously weaving actions and observations to optimise learning (Bramley et al., 2015,0,0; Davis et al., 2020a; Gong et al., 2022). However, the ability to pick informative interventions is not trivially explained. Potential sequences of interventions, i.e. policies, in any given situation are essentially infinite. Yet, this is a problem that humans appear to seamlessly solve throughout their lives.

Previous research has used causal interventions to formalise what constitutes an action (Pearl, 2009). In Pearl's do-calculus, interventions, denoted $do(X = x)$, allow the observer to fix a variable X at a value x, such that $P(X = x) = 1$ while removing links from X's parents. This is done by changing the underlying model during the intervention (Gopnik et al., 2004; Steyvers, 2003). There is ample evidence that participants are sensitive to interventions defined as such and use them to resolve uncertainty and uncover causal structure (Rothe et al., 2018; Lagnado and Sloman, 2002,0,0; Steyvers, 2003; Rottman and Keil, 2012; Bramley et al., 2017a,0; Davis et al., 2020b). Lagnado and Sloman (2004) proposed that participants use their interventions as temporal cues for causal inference. Davis et al. (2018) and Rehder et al. (2022) showed that in cyclical models, carefully targeted interventions enabled learners to temporarily break loops and simplify causal structures. By working in continuous time, but restricting variables to discrete states, Gong et al. (2022) were able to model participants' active learning. Through resource-rational analysis (Lieder and Griffiths, 2019), they showed that participants chose actions by trading off information gain and computational cost. In continuous time and continuously valued space however, accounts are more descriptive and such a formal model has yet to emerge. Bramley et al. (2017b) showed that successful participants in continuous time tended to spread out their interventions in time allowing for periods of observation of the consequences of their actions. Bramley et al. (2018b) further provided a high-level descriptive account of interventions in such a setting, showing that participants' actions in a physics simulation were akin to controlled mini-experiments testing specific hypotheses about object properties. Rehder et al. (2022) showed that participants adaptively changed their intervention strategies to suit the constraints of the experimental context, e.g. intervened for longer when their was a longer delay between a cause and its effect. While these describe features of participants' policies, they do not provide a systematic account of how actions are composed and how such compositions relate or are adapted to the constraints of the task.

Our aim for this project is to propose a systematic qualitatitve analysis of the interventions that participants take in the continuous time and one-dimensional state

[1]Demo: `https://vbtesh.github.io/less_is_more_2023.html`
[2]Data and code: `Github repository`

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

space setting of Ornstein-Uhlenbeck networks with three variables (Davis et al., 2020a; Rehder et al., 2022) to lay the ground work for a formal computational account of action selection. We will attempt to provide a framework to classify interventions and find regularities in their compositions. We wish to understand actions as sequences of lower order discrete moves which are composed together to form complex interventions. Extrapolating from previous work (Bramley et al., 2018b; Rehder et al., 2022), we expect that participants' policies will be tailored to test different hypotheses. Specifically, we expect the composition of interventions to depend on the latent graph structure participants aim to recover.

Ornstein-Uhlenbeck networks OU networks were introduced by Davis et al. (2020a) to study continuous time causal structure learning. They are a collection of interdependent continuous Gauss-Markov processes, which provide a mathematically well behaved framework to model continuous noisy data. Its elegance lies in the fact that causality, through time-dependencies between the variables, is part of its core functional structure. OU networks can be represented as dynamical Bayesian networks, where the variables at a given time t are parents of themselves and other variables at time $t+dt$, where dt is a small time increment. If X, Y and Z are three random variables described by OU processes, the probability of finding the process X at some value x at time $t+dt$ is

$$p(x_{t+dt}|x_t, y_t, z_t) \sim \mathcal{N}\left(x_t + \theta\left(\mu(x_t, y_t, z_t) - x_t\right)dt, \sigma\sqrt{dt}\right) \tag{1}$$

where $\theta > 0$ describes the general strength of all causal effects in the graph, $\sigma > 0$ is the standard deviation of the process over one second, dt is the time step, i.e. the reciprocal of the number of frames per second, $\mu(x_t, y_t, z_t) = x_t + \gamma_{yx}y_t + \gamma_{zx}z_t$ and $\gamma_{yx}, \gamma_{zx} \in \{-1, -0.5, 0, 0.5, 1\}$ represent respectively the causal strength of Y and Z on X. We allow for both weak and strong, and negative and positive links, which is a departure from previous work. See Davis et al. (2020a) for a more comprehensive account.

Methods

We use data from four experiments conducted in the same general framework which we detail below. The principle, borrowed from Davis et al. (2020a), is to provide participants with an interface with three sliders representing three variables and a graph plotting in real time the values of those variables. Variables values are generated from Ornstein-Uhlenbeck processes and depend on one another according to a latent causal graph structure. The task is to observe variations and use the sliders to intervene to recover the causal graph (Figure 1 A.).

Participants Across all four experiments, we recruited 421 participants (264 males, age 26.93 ± 8.41) from Prolific.co, 60 for experiment 1 (36 males, age 26.4 ± 6.66), 121 for experiment 2 (78 males, age 24.58 ± 6.30), 120 for experiment 3 (86 males, age 28.92 ± 9.57) and 120 for experiment 4 (64 males, age 27.8 ± 9.31).

Materials The main interface was the dynamical causal learning game (Figure 1). We let participants intervene on each variable using three sliders. Only one intervention is allowed at a time but they can last for as long as they wish and take, at any point, any value in the range $[-100, 100]$. The main trial interface was designed with the following constraints. First, trials lasted 60 sec at a update rate of 5 Hz, which generated 300 data points. Second, for visual aid, participants were presented with a plot showing 10 seconds of history, interventions were plotted as background shading of the colour of the relevant variable, i.e. *blue*, *green* and *red*. Thus a blue shading on the graph indicates that these data were generated during an intervention on the *blue* variable. Third, we allowed participants to report links during the trial and for an unlimited amount of time after. Finally, we know from Rehder et al. (2022) that higher update rate dt and lower θ can lead to significant differences in performance by rendering causal effects respectively more sluggish or less impactful. Therefore we adopt a form where the update rate $dt = 200$ms which is within the ranges they tested and a $\theta = 0.5$, which is equivalent to their $\omega = \theta dt = 0.1$, which is the value they used in their rigid, i.e. more responsive, condition. The noise parameter σ was set to 3. All parameters were kept constant in all four experiments.

Design and procedure First, to avoid overwhelming participants with mathematical representations, we gave them an illustrated introduction to causality by using informal graphical models. Link strength was expressed simply by showing that given two nodes A and B, one could draw one arrow to indicate a moderate causal relation and two arrows to indicate a stronger one. The direction of the relationship could be specified with negative or positive sign next to the arrow. It was specified that two arrows meant twice as strong to preserve the relative distance between links, therefore all judgements about graphs were collected using sliders with values in $\{-2, -1, 0, 1, 2\}$.

Participants, performed four dynamical learning game trials in the first three experiments and five in the fourth one. The graphs used as ground truths were of two kinds. The first category included generic predefined structures (Figure 1 B.) such as chains, confounded chains or colliders for which variables had neutral labels (*red*, *blue* and *green*). The second category included graphs with

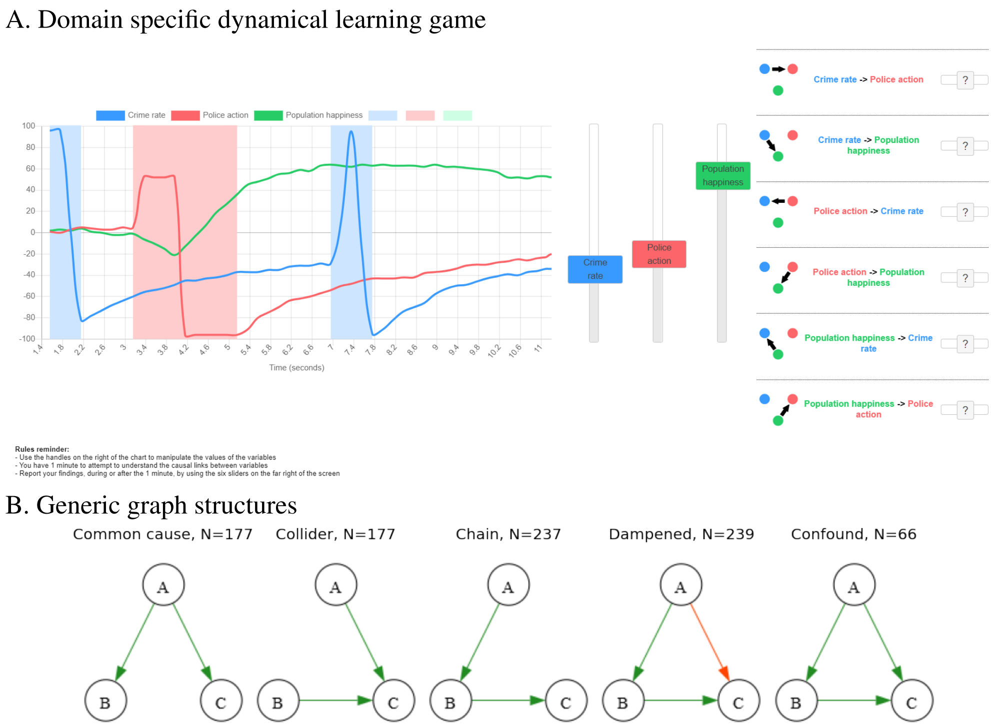

Figure 1: A. Dynamic causal learning game interface. The data is plotted on the graph in real time and the sliders move alone according to the position of the variables in the graph B. Generic causal structures used in the experiments. Green arrow indicate positive relations and red negative ones. The positions and signs were counterbalanced. Other trials using labelled variables made up the rest of trials $N = 875$.

meaningful labels (*Crime Rate*, *Police Action* and *Population Happiness*) and their structure depended on each participant's prior judgement over how those concepts relate, which means that there was no set structure.

Behavioural results

As our focus for this project is on interventions, we pool all the data from all four experiments together, yielding 1804 trials in total for the 421 participants. Figure 2 depicts two trials taken by participants, one for a common cause structure, the other for a chain graph. Overall, participants correctly recovered the exact value of links 59% of the time, significantly higher than the chance level of 20% ($t(420) = 43.1, p < .001$). Unsurprisingly, participants did better with graphs that did not involve indirect effects such as common cause (links recovered: 69%) and collider (70%) and did worse with chains (60%) or dampened (58%) graphs, respectively adding omitting the direct link between the first and last variable in the chain, a bias which has been previously observed in the literature (Fernbach and Sloman, 2009; Bramley et al., 2015; Davis et al., 2020a; Rehder et al., 2022). Participants were excellent at recognising positive ($\mu = .79, t(420) = 39.23, p < .001$) and negative ($\mu = .70, t(420) = 28.71, p < .001$) links compared to chance level (.33). The difference was sig-

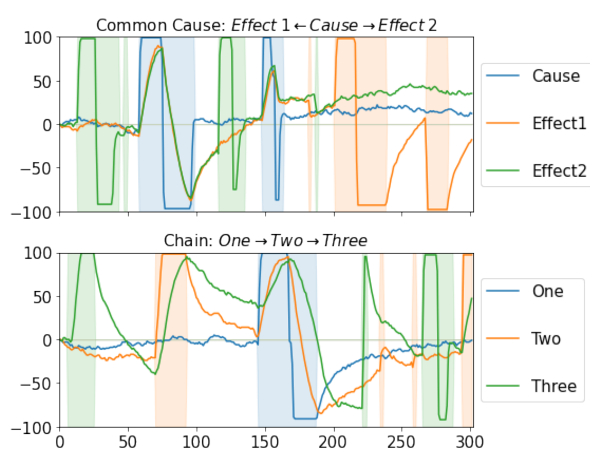

Figure 2: Examples of trials for two participants. The values of variables are on the y-axis and the time point is on the x-axis. Shaded areas indicate interventions and the background colour the variable being intervened upon.

nificant ($\mu = .08, t(840) = 4.81, p < .001$), suggesting that that positive links were easier to recover than negative links. They were less able to differentiate between weak ($\mu = .40, t(420) = 5.23, p < .001$) and strong ($\mu = .43, t(420) = 6.38, p < .001$) links compared to chance level (.33). The difference was between weak and strong links was not significant. Overall, participants' ability to recover the correct graph raises the question of how they did it; more specifically what types of behaviour, observations and interventions, did they performed to achieve such performances?

Analysis of interventions

Participants made heavy use of interventions, spending on average 43% ($\pm$29%) of their time intervening during each trial. While there may be within participants consistencies and strategies, for this short work, we pool all interventions together to provide a global overview. We define an intervention as any uninterrupted sequence of frames in a trial where a slider is clicked. Across all experiments, we recorded 14445 interventions. Participants made on average 8.43 ($\pm$4.37) interventions per trial for a mean length of 3.12 ($\pm$3.40) seconds. To analyse these interventions and extract patterns of behaviour, we propose to first provide quantitative indicators describing single interventions, use these measures to classify interventions in different types and finally use this classification to look for differences between participants and underlying ground truth structures.

Describing single interventions using summary statistics

In general, because time (5Hz) and the space of variable values ($[-100, 100]$) are pseudo continuous, the number of possible interventions is so large that every intervention recorded is unique. To find consistencies for further analysis, we propose to describe interventions from a set of summary statistics which we content encode some representation of participants' intentions when intervening. Figure 2 and Figure 4 provide instances of recorded interventions. We note two key features of these actions. First, the emphasis on "holding", in that participants seem to hold a variable at or around a specific value for a significant proportion of the action, suggesting that a key statistic of any intervention is its mode, i.e. the value most often taken during an intervention. Second, the presence of "swipes", i.e. movements from one end of the range to the other at a value which is close to the previous one, only with a reversed sign. This suggests that participants tend to peg variables at the bounds of the range, irrespective of the sign. Simply put, participants treat pulling down or pushing up, given that the initial value of the variable is about the origin, as equivalent. We argue that any intervention is essentially a composition of swipes and holds. At the extremes, many swipes with little or no holding in a short amount of time yields "wiggling" behaviour and few or no swipes with constant holding yields simple pegging of a variable.

Given this initial qualitative outlook, we propose a set of four summary statistics: *length in seconds*, the duration of the intervention, represented in the log space as the length of interventions followed a log normal distribution, *mean holding absolute value*, the mean absolute value the variable took while the participant was holding, *holding time (in percent)* the amount of time spent holding as a percentage of the total length of the intervention and *number of swipes* the number of swipes in the intervention, i.e. how many times the direction of the intervention path changed.

Classifying interventions

To better understand the types of interventions participants make, we propose to use the statistics discussed above to classify interventions. We fit a variational Bayesian estimation of a Gaussian mixture (Blei and Jordan, 2004) using Python's scikit-learn library. We fit a model with 6 components and diagonal covariance matrices for each of them. To pick the number of components, we notice that for 4 and 5 components, the mixture distribution always assigns 63% of all interventions to a single component in which interventions are distributed similarly to the full sample. This means that the model tends assign the bulk of all interventions to a single component. At 6 however, fitting the model yields a balanced mixing distribution over the components with each of them having distinct distributions over the four measures. We do not claim that those clusters represent a segmentation of the space from which participants sample their interventions. More humbly, our aim with this classification is to extract structure from the dataset and test whether thinking about interventions as compositions of holds and swipes is a relevant approach.

We categorise each intervention by assigning it to the component maximised by the model's posterior. We name the six clusters as follows according to their mean on the four measures used in the mixture: *behavioural noise* (6.5%), simple short clicks about the origin, *single swipes* (36.7%), *two swipes and medium holding time* (8.9%), *single swipe and long holding time* (24.4%), *two swipes and long holding time* (21.5%) and finally *more than two swipes and medium holding time* (1.9%).

Figure 3 provides an overview of the full set of interventions. From plot A. we can see that most interventions lie between one and three swipes with a holding time of about 50%. Plot B.1. depicts the frequencies for each cluster. As expected, interventions classified as behavioural noise tend to be less numerous than other types and from plot B.2. and are also the least informative. The

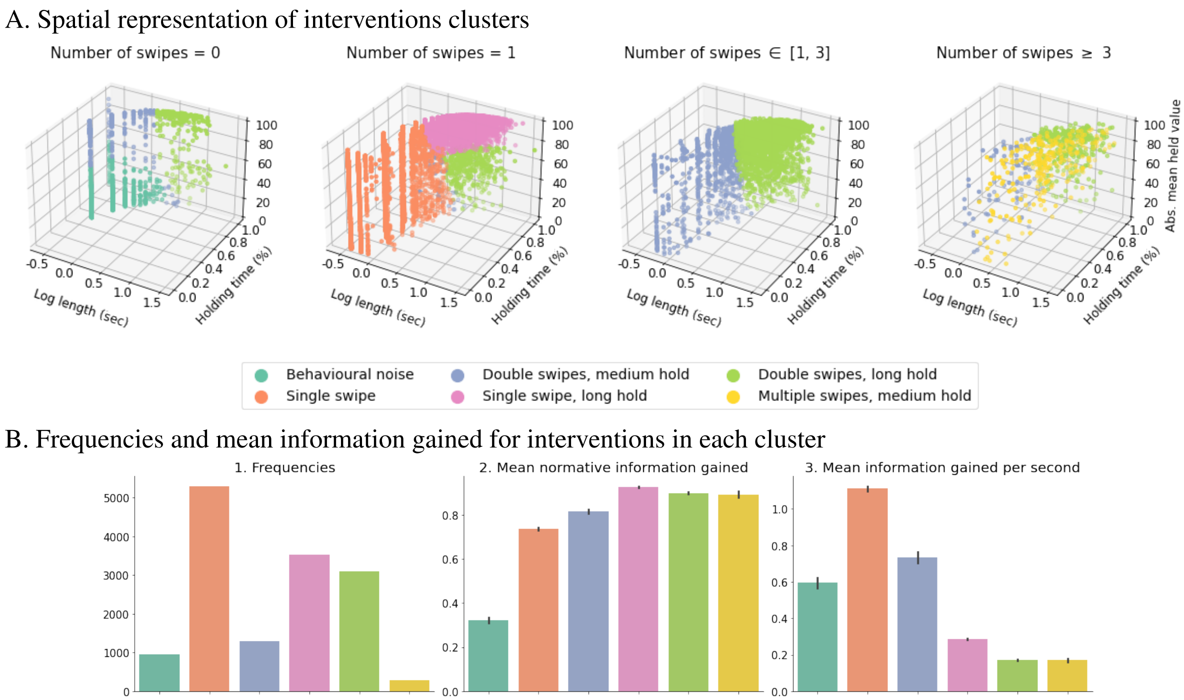

Figure 3: A. Spatial representation of the clusters. Intervention length, time spent holding and the absolute mean held value are plotted in 3d and then split between the number of swipes. B. Summary statistics for each cluster. 1. Number of interventions in each clusters, order in ascending order of length 2. Mean information gained from performing interventions in each cluster when starting from a uniform prior over link values 3. Mean information gained per second interventions in each cluster.

most common interventions are single swipes, which, from plot B.3. are the most information efficient type of interventions, meaning they yield on average the most information about the graph in the shortest amount of time. This is because swiping is normatively more informative than holding as it generates larger value changes, putting interventions with a long holding time at a disadvantage. The second and third most regularly performed interventions are single swipes with long holding time and double swipes with long holding time. These are the most informative but are less time efficient. The fact that participants favour holding suggests, as expected, that the way they extract information is not normative. Finally, interventions with multiple swipes and shorter holding time, i.e. wiggles, are rare.

Intervention choices depend of the ground truth causal structure

We ask whether participants' interventions depend on the structure of the underlying graph they attempt to retrieve. We split the different cases of causal structures in three qualitatively distinct categories, ignoring link sign and strength and focusing on the detection of the direct outgoing links from the held variable. Thus, for each node on which to intervene, there can be three cases of underlying structure: two outgoing links, one outgoing link, or no outgoing links. Across all studies, interventions appear to highly depend on the structure of the underlying graph ($\mathcal{X}^2(10, N = 14445) = 486.1, p < .001$). While interventions classified as noise seem to mildly depend on graph structure ($\mathcal{X}^2(2, N = 945) = 7.07, p < .05$), interventions consisting of a single swipe and long holding time ($\mathcal{X}^2(2, N = 3533) = 232.2, p < .001$) were performed about four times more often in cases with two outgoing links than they were in cases with no outgoing links. This suggests that holding is instrumental in determining the strength and direction of a link. The other distinctive difference was that despite being rare, the overwhelming majority of wiggles, i.e. interventions with many swipes, were performed in cases with no outgoing links, where they were present twice as often as in other cases ($\mathcal{X}^2(2, N = 281) = 122.9, p < .001$), suggesting that wiggling is favoured for disambiguating whether there is a link or not but not as effective at characterising its sign and strength.

Swiping serves causal discovery and holding allows estimation of link direction and strength

The proposition that swiping is more effective to determine the absence of a link and that holding serves more to qualify its direction and strength is further corroborated by the fact that their respective use frequen-

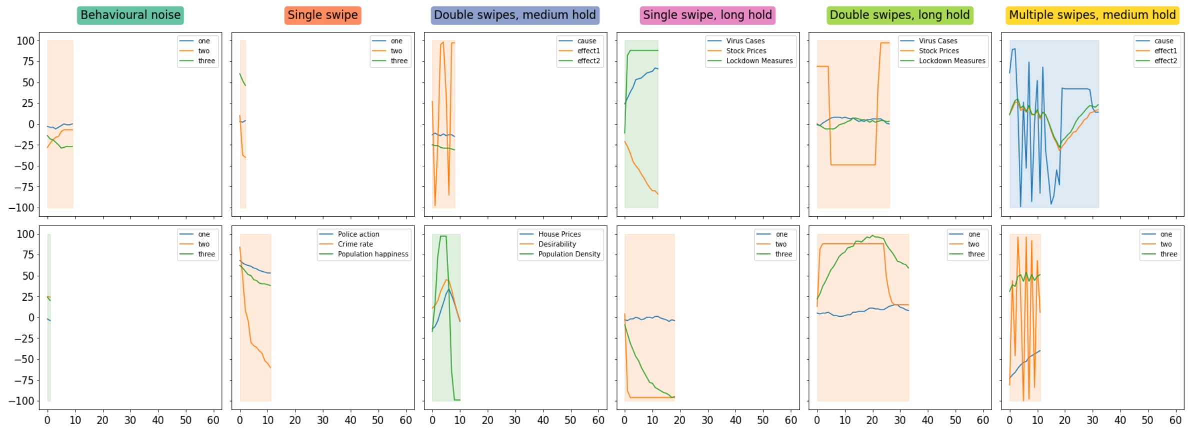

Figure 4: Sample interventions from each of the clusters. Each column contains samples from one intervention type, the x-axis is the length in frame and is shared among all samples to illustrate the different lengths across clusters.

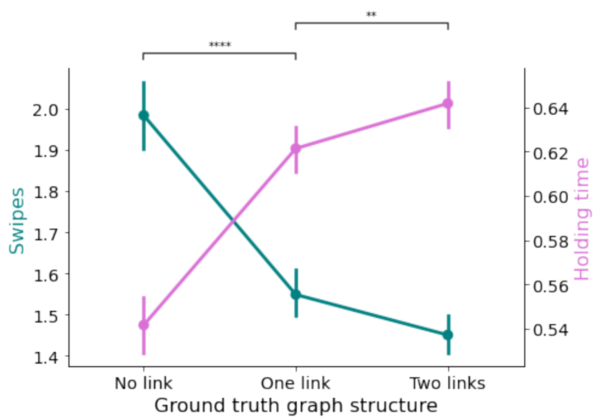

Figure 5: Mean number of swipes performed on a node and mean holding time of each action performed on a node for each case of underlying structure: no link, one link and two outgoing links. All pairwise differences were significant.

cies depended directly on the underlying graph structure, i.e. whether there is no, one or two outgoing links from the held node. Indeed, both the number of swipes ($F(5409) = 77.74, p < .001$) and the time spent holding ($F(5409) = 77.51, p < .001$) a node depended highly on graph structure. Figure 5, shows that this pattern is symmetrical and that the more links there are, the less swipes are performed and the longer nodes are held fixed. This suggests that detecting a link is decoupled from qualifying its strength and direction. For the former, swipes are more effective, generating rapid observable change and thus evidence for a connection, while for the latter, holding a node and observing the nature of the effect in the other two is more efficient. This is a strong departure from standard Bayesian agents which can do both at the same time. Instead, humans seem to need to decompose the inference process in two distinct phases.

General discussion

The fact that intervention policies strongly depend on underlying graph structures provides guiding evidence for future research to uncover participants' inference models. The tendency to use multiple swipes to validate the absence of a link for instance may suggest that participants are engaging in some form of counterfactual reasoning following their initial swipe, e.g. perhaps the data generated by the initial swipe would have happened anyway, or alternatively perhaps the lack of change was the result of underlying indirect dynamics, therefore justifying additional swipes. Furthermore, the need to hold a variable still to infer the direction and strength of a link indicates that participants require the cause to be stable and predictable in time in order to accurately evaluate the nature of its effects. It favours models that predict future changes in the candidates effects given a fixed value of the held node. This disfavours models which simply use Bayesian updates on the observed data such as the Local Computation model (Davis et al., 2020b). The more recent Causal Event Abstraction model (Rehder et al., 2022), by being predicated on the idea that participants need to hold to collect evidence for the presence of a link seems like a better candidate. However, it cannot explain the abundance of swipes in the absence of a link. Both remain accounts of learning as simple evidence accumulation processes and stay silent about the dynamics of how actions are adapted online to the generated data. Future modelling efforts should propose dual models of observations and actions by building on the present findings of interventions as compositions of swipes and holds and an inference process which satisfy the constraints of participants' action policies.

References

Blei, D. M. and Jordan, M. I. (2004). Variational inference for Dirichlet process mixture models Mean field variational inference. *Bayesian Analysis*, (1):1–9.

Bramley, N. R., Dayan, P., Griffiths, T. L., and Lagnado, D. A. (2017a). *Formalizing neurath's ship: Approximate algorithms for online causal learning*, volume 124.

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

Bramley, N. R., Gerstenberg, T., Tenenbaum, J. B., and Gureckis, T. M. (2018b). Intuitive experimentation in the physical world. *Cognitive Psychology*, 105(May):9–38.

Bramley, N. R., Lagnado, D. A., and Speekenbrink, M. (2015). Conservative Forgetful Scholars : How People Learn Causal Structure Through Sequences of Interventions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(3):708–731.

Bramley, N. R., Mayrhofer, R., Gerstenberg, T., and Lagnado, D. A. (2017b). Causal learning from interventions and dynamics in continuous time. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, volume 1, pages 150–155.

Davis, Z. J., Bramley, N. R., and Rehder, B. (2020a). Causal Structure Learning in Continuous Systems. *Frontiers in Psychology*, 11:1–29.

Davis, Z. J., Bramley, N. R., and Rehder, R. (2020b). The paradox of time in dynamic causal systems. *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society*, pages 808–814.

Davis, Z. J., Bramley, N. R., Rehder, R. E., and Gureckis, T. M. (2018). A Causal Model Approach to Dynamic Control. *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*, pages 281–286.

Fernbach, P. M. and Sloman, S. A. (2009). Causal Learning With Local Computations. *Journal of Experimental Psychology: Learning Memory and Cognition*, 35(3):678–693.

Gong, T., Gerstenberg, T., Mayrhofer, R., and Bramley, N. R. (2022). Active causal structure learning in continuous time. *arXiv*.

Gopnik, A., Sobel, D. M., Danks, D., Glymour, C., Schulz, L. E., and Kushnir, T. (2004). A Theory of Causal Learning in Children: Causal Maps and Bayes Nets. *Psychological Review*, 111(1):3–32.

Lagnado, D. A. and Sloman, S. A. (2002). Learning Causal Structure. In *Proceedings of the Annual Meeting of the Cognitive Science behaviour*, pages 276–281.

Lagnado, D. A. and Sloman, S. A. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning Memory and Cognition*, 30(4):856–876.

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

Lieder, F. and Griffiths, T. L. (2019). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*.

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

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, 3(September):96–146.

Rehder, B., Davis, Z. J., and Bramley, N. (2022). The Paradox of Time in Dynamic Causal Systems. *Entropy*, 24(7):863.

Rothe, A., Deverett, B., Mayrhofer, R., and Kemp, C. (2018). Successful structure learning from observational data. *Cognition*, 179(March 2017):266–297.

Rottman, B. M. and Keil, F. C. (2012). Causal structure learning over time: observations and interventions. *Cognitive psychology*, 64(1-2):93–125.

Sobel, D. M., Tenenbaum, J. B., and Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science*, 28(3):303–333.

Steyvers, M. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, 27(3):453–489.