

Causal reasoning in a prediction task with hidden causes

Pedro A. Ortega (ope@seas.upenn.edu)

School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, PA 19104 USA

Daniel D. Lee (ddlee@seas.upenn.edu)

School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, PA 19104 USA

Alan A. Stocker (astocker@sas.upenn.edu)

Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract

Correctly assessing the consequences of events is essential for a successful interaction with the world. It not only requires a causal understanding of the world but also the ability to distinguish whether a given event is the result of an agent's own action (intervention) or simply the consequence of the world being in action (observation). Previous studies have shown that humans can learn causal structures, and that they can distinguish interventions from observations. These studies almost exclusively focused on structures where interventions led to a simple forward conditioned inference problem. We tested human subjects in a prediction game that required the integration over hidden causes, using a betting mechanism that allowed us to monitor subjects' beliefs. Subjects learned the causal structure and the conditional probabilities with appropriate feedback. Once learned, all but one were immediately able to correctly predict the causal effects of their interventions according to optimal causal reasoning.

Keywords: Causal interventions, betting game, belief updates.

Introduction

There is ample experimental evidence suggesting that humans guide their decisions in interacting with the world using causal knowledge. Causal interpretations offer superior explanatory power over purely observational (correlational) ones, allowing us to understand, predict and interact with the world in a multitude of ways. Indeed, the processing of causal information appears to be deeply embedded in animal cognition (Sloman, 2005; Blaisdell et al., 2006); and studies in child development have shown that we learn to form and rely on causal knowledge early on in our lives (Gopnik et al., 2004; Meltzoff, 2007).

Hagmayer and Sloman (2009) have recently proposed the *causal theory of choice*. The theory proposes that humans correctly infer the consequences of their own actions (interventions) based on causal models that they have learned through experience. Crucially, it predicts that reasoning based on an intervention vs. an observation generally results in different beliefs about their consequences even if the observation and the intervention are identical. The theory uses a popular probabilistic framework called *causal Bayes nets* that permits a precise mathematical formulation of causal reasoning under this distinction (Spirtes and Scheines, 2001; Pearl, 2009; Dawid, 2007).

This theory has been put to test in several studies. In one-shot decision making, it has been shown that humans treat their own decisions as interventions according to a causal model (Hagmayer and Sloman, 2009), and that they have

different beliefs if the decisions were simply observed instead (Saito and Shimazaki, 2011). Subjects in these experiments were given a written description and a graphical representation of the causal structure of the task. Crucially, however, they did not experience the consequences of their decisions, indicating that their differentiation between intervention and observation is inherent and spontaneous rather than learned. Subsequent studies have then investigated the degree to which human observers can learn the causal structure over repeated observations of the consequences of their interventions (e.g., Steyvers et al. (2003)). Recently, Hagmayer and Meder (2013) studied human subjects in a repeated decision-making task where the subjects' goal was to maximize payoff by intervening in a causal system with multiple outcome variables. By introducing abrupt changes to the causal structure of the system that kept the observable consequences constant, the experimenters were able to elicit revisions in the subjects' decisions, thereby showing that they were indeed sensitive to the causal structure.

However, these previous studies have been limited to causal structures in which interventions led to a simple forward conditioned inference problem where the outcome was only conditioned on the intervention itself (e.g., "causal chain" or "common cause" problem; see Fig. 1a,b). Thus it remains unclear whether human subjects can learn and correctly access the consequences of their intervention in more complex causal structures (see e.g., Berry and Broadbent (1995); Osman (2010)). An exception is the work by Meder et al. (2009), in which participants were asked to report the probabilities of events that required them to combine observational/interventional evidence and the base rates of a hidden cause. In this study, however, subjects were explicitly informed about the underlying causal structure at the beginning of the experiment.

The main goal of our work was to investigate whether humans subjects can learn more complex causal structures, in which their interventions only led to a partial conditioning of the outcome. Specifically, we tested subjects in a repeated prediction task that required them to marginalize over the hidden cause in a "fully connected common cause" structure (Fig. 1c). Our hypothesis was that it is sufficient to let subjects experience both the observational and interventional regimes of a controllable variable in a sequence of events for them to learn an accurate causal model of this structure. We

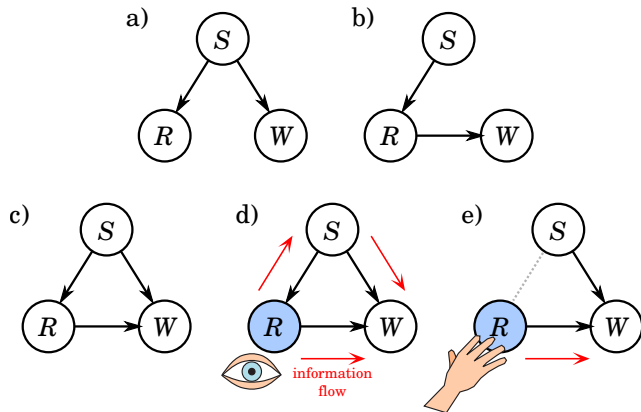


Figure 1: Intervention *vs.* observation in a directed acyclic graph (DAG) over the binary random variables S , R and W . The DAGs of two causal models known as a) “common cause” and b) “causal chain” that have been previously studied. c) The DAG of a more complex, “fully connected common source” structure. d) Observing the value of R entails a belief update captured through Bayesian conditioning. e) In contrast, setting the value of R (intervention) leads to a modified graph, in which R is independent from S . As a result, the information flow between R and W is different compared to the observational regime in panel (d), leading to a different Bayesian belief update. Crucially, compared to the simpler structures that have been previously tested (*e.g.* causal chain), correct inference requires integration over the hidden cause S .

implemented the task in form of a color prediction betting game. We designed a clever betting mechanism that directly reflects subjects’ beliefs on a trial-to-trial basis based on a penalty function that is minimized when reporting the true beliefs (Dawid, 2006). This allowed us to track subjects’ belief updates and thus to quantitatively confirm the degree to which subjects conform to causal Bayesian reasoning. We found that four out of five participants were able to precisely learn the form of this more complex causal structure and its associated beliefs when provided with sufficient evidence. Furthermore, once learned, subjects were immediately able to correctly apply their knowledge of the causal structure and the associated beliefs when performing interventions in situations where the common cause was hidden.

Inference with Causal Interventions

The distinguishing feature of a causal intervention is that it renders the causally preceding random variables statistically independent from the intervened random variable. Colloquially, this means that “actions can influence the future but they cannot amend the past”. These causal relations can be formalized in terms of a directed acyclic graph (DAG) (Pearl, 2009). Consider a random variable R embedded in a causal context given by S and W representing parent and child random events as depicted in Fig. 1c. This fully connected “common cause” graph is the simplest model where the effect is

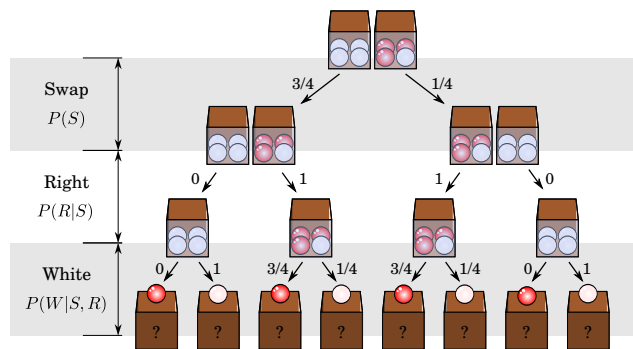


Figure 2: The color prediction betting game. Each trial of the game followed three steps: (S) The computer swapped the position of two boxes with probability $P(S)$. (R) The computer chose the right box with probability $P(R|S)$. (W) Finally, the computer drew a white ball from the chosen box with probability $P(W|S,R)$. Subjects had to place a bet on the color of the drawn ball before the color was revealed. In the Training Game 2 and the Test Game, subjects sometimes were asked to intervene at the middle step (R).

controlled by a single unknown cause S through two different pathways in which one of them can be intervened.

The probability of W given the observed value of R (Fig. 1d) can be computed as

$$P(W|R) = \frac{\sum_s P(W|S=s,R)P(R|S=s)P(S=s)}{\sum_s P(R|S=s)P(S=s)}. \quad (1)$$

If, however, the value of R is the result of an intervention ($\text{do}(R)$), then

$$P(W|\text{do}(R)) = \sum_s P(W|S=s,R)P(S=s) \quad (2)$$

is the predicted probability of W (Pearl, 2009). Note that (2) is obtained by computing the posterior on the modified DAG (Fig. 1e).

Experimental Method

The general goal was to experimentally test whether subjects can update their beliefs in a way that is consistent with causal interventions when they choose the values of random variables themselves. We engaged subjects in a sequential betting game where they repeatedly had to bet on the outcome of a random event. One trial of the game comprised three sequential steps, each one associated with a different random variable; and where subjects sometimes were asked to choose the value of the middle variable themselves (intervention). The rationale behind this sequential setup is twofold. First, we wanted to succinctly capture the full scope of an intervention, and this setup allowed us to measure the effect on both causally preceding and succeeding random variables. Second, we wanted to test whether humans can dynamically switch between their belief updates for an observation and the

updates for an intervention. Third, we wanted to continuously monitor subjects learning rate throughout the entire experiment. To the best of our knowledge, these features combined are unique to our experimental design.

The experiment was organized as a set of two simplified versions of a *color prediction betting game* intended to train the subjects, followed by the main game that tested our hypothesis (Fig. 2). The goal of the first game was to familiarize subjects with the betting mechanism, while the second game was aimed to let subjects learn the causal structure linking the random events. Finally, the main game tested their causal reasoning skills. Each game was subdivided into a sequence of *levels* (blocks) containing ten prediction trials which are described further down.

Subjects were instructed to bet on the outcome of the last random variable. The game structure was such that they were highly motivated to maximize their wins X by making accurate bets. The wins were determined on a level-to-level basis as follows: At the beginning of each level, subjects were given a fixed budget B which they had to protect from losses L that occurred in each trial due to the inaccuracies in their bets. If they passed the ten trials of a level with a positive budget $B - L > 0$, they collected the remaining points and added them to their total wins as $X \leftarrow X + (B - L)$. If, however, the budget turned negative at any moment before the end of a level, subject had to repeat the level from the beginning. This design encouraged subjects to improve their predicting strategies early on in order to successfully progress in the game.

Game structure. We designed a betting game where subjects were required to bet on the color of a ball drawn from one of two boxes. The structure of one trial of the game is illustrated in Fig. 2. Two boxes were initially placed next to each other, the one on the left containing four white balls and the one on the right containing three red and one white ball. In the first step, the positions of the boxes were swapped with probability $p = 1/4$. In the middle step, one of the two boxes was chosen, either by the computer or by themselves. In the final step, a ball was randomly drawn from the chosen box and its color revealed. The game is formalized *via* three binary random variables S, R and W whose values were drawn from the (conditional) probability distributions $P(S)$, $P(R|S)$ and $P(W|S, R)$ respectively. The corresponding causal DAG is shown in Fig. 1c. Importantly, the computer always chooses the box with more red balls in the middle step, which is something the subjects could learn from data.

Betting process. On each trial, subjects were forced to place a bet on the color of the drawn ball W before its actual color was revealed and the bets were processed. The accuracy of their bets was measured using a *log-loss scoring rule* (Dawid, 2006; Bickel, 2007). Specifically, subjects made their bets by indicating the losses they were willing to accept for each of the two outcomes. They did so by adjusting the lengths of two coupled bars (see Fig. 3).

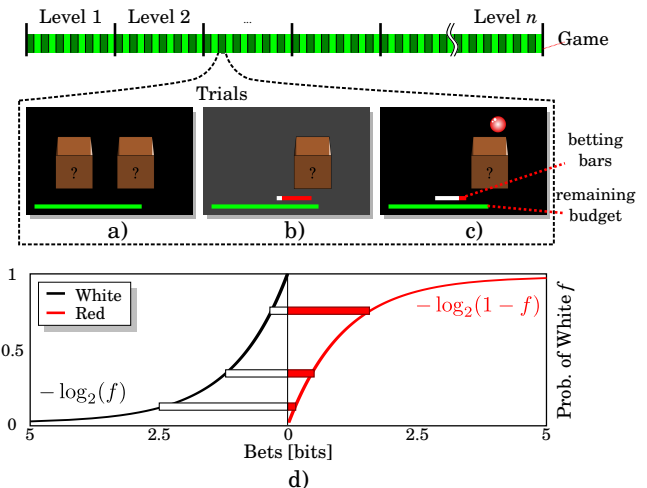


Figure 3: Structure and screenshots from a trial in the Test game. a) At the beginning of the trial, the boxes were placed next to each other and swapped with a given probability. In contrast to the two Training games, the boxes were opaque and thus did not show their content. b) After one of the boxes (here: the right one) was selected with a certain probability, subjects placed a bet by adjusting the lengths of two coupled betting bars (red/white). The length of each bar determined the amount the subject was willing to lose from the budget (in green) in the event of each outcome (*i.e.*, the ball being “white” or “red”). c) Once the bet was placed, a ball was drawn randomly and its color was revealed. The betted amount got subtracted from the budget. d) The betting structure was such that subjects who minimized their expected losses were implicitly reporting their beliefs (see text for details).

The lengths were calculated as $L(\text{white}) = -\log_2(f)$ and $L(\text{red}) = -\log_2(1-f)$ respectively, where $f \in (0, 1)$ stands for the subjects’ predicted probability for white. The advantage of this scoring rule is that it is *strictly proper*: theoretically, a strictly proper scoring rule is uniquely optimized by the true probabilities, thus encouraging subjects to quickly adopt the true probabilities and be honest in reporting their beliefs, while simultaneously allowing us to accurately measure their beliefs on a trial-per-trial basis rather than estimating them from empirical averages.

The bets were processed as follows. Once the ball was drawn and the color has been revealed, the corresponding loss was subtracted from the subject’s budget. To prevent subjects from being successful through conservative, indifferent guessing (*i.e.*, choosing $f \approx 1/2$ with an associated loss of $L \geq 1$ bit), they were allocated an initial budget of only $B = 10$ bits per level. Additionally, the log-loss renders confident guesses too risky, as penalties for incorrect guesses diverge rapidly when $f \rightarrow 0$ or $f \rightarrow 1$ (see Fig. 3d). While this betting scheme may appear rather complicated, subjects quickly assimilated it within the first few trials after which it

reportedly felt very natural.

Game sequence. Subjects played a set of three games in sequence, namely two Training games and one Test game. All games followed the general description above and had the same probability structure. They differed in three aspects: i) the number of levels subjects had to complete successfully, ii) the transparency of the boxes, and iii) whether subjects were asked to intervene. With transparent boxes, subjects can form their beliefs for “white” by simply looking at the contents of the chosen box. However, if the boxes are opaque, then subjects must infer the contents by combining prior knowledge about S and the known value of R . This is a much harder prediction task because subjects need to marginalize over two possible states of S in order to be able to bet accurately. Also, when asked to intervene subjects can attempt to choose the box containing only white balls in order to simplify the prediction task. Obviously, they cannot pick this box with certainty when the boxes are not transparent. Below is a short description of the three games and their differences:

- *Training Game 1:* The main purpose of the first game was to familiarize subjects with the task and the betting mechanism that is based on the log-loss scoring rule. Subjects passively observed the random events generated by the computer and then placed their bets before the ball was drawn from the chosen box. The contents of the boxes were visible at all times.
- *Training Game 2:* The purpose of this game was to permit subjects to learn the causal structure of the game. The game was similar to the first training game in that the contents of the boxes were fully visible at all times. However, subjects were asked to choose the box in 50% of the trials. Crucially though, since the boxes are transparent, the subjects’ bet on the color of the ball W depended only on the values of S and R and *not* on whether they picked the box themselves. Formally, this means that $P(W|S, R) = P(W|S, \text{do}(R))$.
- *Test Game:* The main game tested whether subjects successfully used their causal knowledge acquired throughout the two training games. It was a more difficult game for the subjects because the boxes were no longer transparent. The game allowed us to test whether subjects treated interventions and observations differently when computing their belief updates, and thus whether they were able to marginalize over the latent variable S . As in Training game 2, subjects were asked to pick the box in 50% of the trials. Here, however, the optimal belief updates were different depending on whether they chose the box themselves or not, *i.e.*, $P(W|R) \neq P(W|\text{do}(R))$.

The difficulty of the three games progressively increased. See Table 1 for a summary of their differences.

Table 1: *Game Parameters*

Game	Levels	Transparent	Intervention
Training 1	10	yes	no
Training 2	10	yes	yes (50%)
Test	40	no	yes (50%)

Data Collection. Five students (S1-S5) from the University of Pennsylvania took part in this study after giving their informed consent. All subjects were naïve to the task. This study did not attempt to characterize across population differences.

The games were implemented as computer games (see Fig. 3a,b,c for actual screen shots) and subjects played the three games in sequence on a laptop computer (Lenovo Thinkpad X201). They were instructed to maximize their overall wins X in each one of the games. All subjects completed the three games within less than 90 minutes. Since this required completing at least 600 trials, the average time spent on each trial did not exceed 9 seconds.

Subjects were not given the correct probabilities nor the causal structure of the game, and thus they had to learn these parameters from actually playing the games. However, they were told that all three games used identical statistics, and only varied in whether interventions were allowed or not, and in the transparency of the boxes. Subjects were paid \$10 (in U.S. Dollars) for their participation and an extra \$10 if they completed all three games (which they did without exception).

Results

Figure 4 summarizes our main findings. It shows the empirical frequencies of the different conditions in the Test game, as well as subjects’ average beliefs about drawing a white ball under each of these conditions. For comparison, we included the optimal beliefs derived from the true causal model (OPT). All subjects approximately learned the correct probabilities associated with choosing the left or the right box (intervention conditions in Fig. 4b). More importantly, with the exception of one subject (S3), all subjects clearly distinguished between interventions and observations as indicated by their highly significant difference between the beliefs assigned to the condition $R = 0$ (observation) as compared to the condition $\text{do}(R = 0)$ (intervention). The measured beliefs for the intervention condition $\text{do}(R = 1)$ are not very informative because subjects tested this intervention only a very small number of times (Fig. 4a), and thus are not shown. In fact, from a point of view of maximizing utility the intervention $\text{do}(R = 1)$ should never be performed. In general, subjects’ behavior is qualitatively well captured by the optimal causal model.

Subjects who distinguished between the interventional and observational regimes were able to make this distinction right from the start. This is reflected in the fact that their beliefs changed very little during the Test game (see Fig. 4c). This

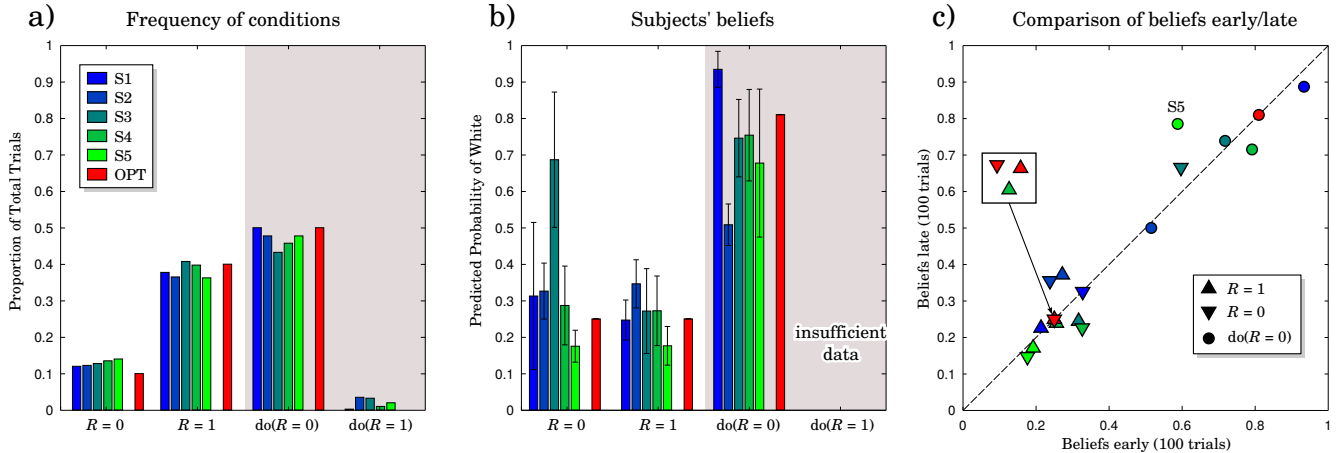


Figure 4: Measured empirical frequencies and average beliefs in the Test game. Data for each individual subject are shown together with the optimal behavior. a) Empirical relative frequencies of the different conditions. Note that subjects for the large part correctly avoided to perform the intervention $\text{do}(R = 1)$. b) Subjects’ average beliefs of drawing a white ball for each condition (error bars represent the standard deviation). The gray shaded areas in a) and b) indicate the conditions when subjects performed interventions. Note, because subjects very rarely chose the intervention $\text{do}(R = 1)$, the resulting empirical belief values were not informative. c) Comparison of the average beliefs during the first and the last hundred trials for each subject and each condition. Points on the diagonal denote no change.

suggests that subjects can directly apply the causal structure learned during the Training games to the new situation of the Test game rather than treating the lack of knowledge of S as a novel context that requires to learn the task’s statistics anew.

In order to analyze the learning behavior of subjects, we compared their trial-by-trial performance with that of an optimal strategy that consistently chose the best bet. The performance is measured in terms of the *cumulative regret* defined as (Bubeck and Cesa-Bianchi, 2012)

$$R(T) = \sum_{t=1}^T L(f_t, w_t) - \sum_{t=1}^T L(f_t^*, w_t) \quad (3)$$

where f_t and f_t^* are the probabilities of drawing a white ball in trial t issued by the subject and the optimal performer respectively, $w_t \in \{0, 1\}$ encodes the actual outcome at time t , and L is the log-loss

$$L(f, w) = \begin{cases} -\log_2(f) & \text{if } w = 1, \\ -\log_2(1 - f) & \text{if } w = 0. \end{cases} \quad (4)$$

The results are shown in Fig. 5. Learning is generally reflected in a negative curvature of the cumulative regret. Notice that intervals having zero slope correspond to optimal behavior where subject’s bets matched the ones of the optimal performer (OPT). Furthermore, super-optimal behavior (*i.e.*, decreasing cumulative regret) can occur in individual realizations of the experiment, but not on average over repeated runs. For example, subject S2 got lucky in the first Training game, by initially performing very risky bets that, by luck, turned out to be correct.

The regret curves shown in Fig. 5 provide some interesting insights. First, they reveal that subjects seem to have learned to correctly use the betting mechanism early on (typically within 40 trials) during the two Training games. Aside from some occasional deviations from the optimal betting strategy that can be attributed to intermittent exploratory and/or risky behavior, the regret curves are surprisingly flat over relative large intervals. In the Test game however, optimal behavior was the exception rather than the norm, with most of the regret curves showing constant positive slopes with very little, if any, curvature. Given that subjects’ average beliefs are remarkably close to the optimal causal model (Fig. 4b), and that these beliefs were already instantiated/learned at the time they started the Test game (Fig. 4c), this suggest that some additional source of uncertainty caused this suboptimal behavior. A simple explanation could be that during the Test game, when the boxes were no longer transparent and the subjects had to remember the correct belief structure of the causal model, uncertainty (noise) in these remembered beliefs led to temporal fluctuations in the subjects’ beliefs (sampling) that resulted in such suboptimal behavior (Gifford et al., 2014). Note, however, that the notion of a “shaky hand”, *i.e.*, noise in adjusting the betting bar, is an unlikely alternative explanation because the regret curves of the two Training Games clearly indicate that subjects were perfectly able to accurately operate the betting mechanism.

Discussion

With the exception of S3, all the subjects made bets that were contingent on whether they chose the box themselves or not. These bets were qualitatively consistent with predic-

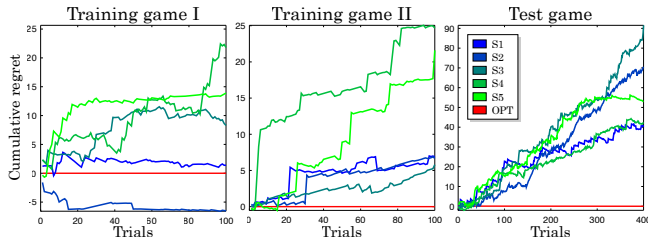


Figure 5: Cumulative regret curves for measuring trial-by-trial learning behavior, shown for all three games. A negative curvature indicates learning, and a slope equal to zero corresponds to optimal behavior. The green-to-blue curves show the prediction performance for individual subjects. The red curve corresponds to the optimal performer.

tions derived from the optimal causal model, implying that they marginalized over the latent cause and treated their own actions as causal interventions in belief updates.

Crucially, this distinction is made immediately at the beginning of the Test Game. Figure 4c, together with the regret curves, suggests that all subjects (with the exception of subject S5) relied on their knowledge learned during the Training Games, but no longer learned during the Test Game. This is interesting, because the task during the the Training games did *not* depend on whether the regime was interventional or not, but only on the contents of the chosen box. Indeed, S3—the only subject who did not learn to distinguish between intervention and observation—actually outperformed the other subjects during the second Training game. Thus, this shows that subjects tend to learn the causal structure of a system and use this knowledge even under conditions where it is not beneficial with regard to expected utility.

Conclusions

Our results complement previous studies that investigated the role of causal reasoning in decision making. We have found that subjects can learn complex causal dependencies in a repeated betting task without being explicitly informed about the underlying causal structure. Furthermore, most subjects in our study performed in a way that optimally combined Bayesian and causal reasoning: first by integrating over latent causes and second by distinguishing between actions and observations when interpreting the feedback signal.

An important open question is under what conditions humans acquire causal knowledge. For this, it is necessary to address the question of what constitutes causal evidence by analyzing how human subjects combine evidence when they can both observe and intervene. In our setup, we let subjects: a) see the events that causally precede their decisions (*i.e.* the underlying cause); and b) experience the consequences of these decisions under the observational and interventional regimes. We have shown that this information was sufficient for all but one subjects. Extending the study to a larger subject pool will provide a more accurate description

of differences across the population.

Acknowledgments We thank the anonymous reviewers for their very valuable comments and suggestions that helped improve the paper. This study was funded by grants from the U.S. National Science Foundation, Office of Naval Research and Department of Transportation.

References

- Berry, D. and Broadbent, D. (1995). Implicit learning in the control of complex systems. In Frensch, P. and Funke, J., editors, *Complex Problem Solving: The European Perspective*, pages 131–150. Hillsdale, NJ: LEA.
- Bickel, J. (2007). Some Comparisons among Quadratic, Spherical, and Logarithmic Scoring Rules. *Decision Analysis*, 4(2):49–65.
- Blaisdell, A. P., Sawa, K., Leising, K. J., and Waldmann, M. R. (2006). Causal Reasoning in Rats. *Science*, 311(5763):1020–1022.
- Bubeck, S. and Cesa-Bianchi (2012). Regret Analysis of Stochastic and Nonstochastic Multi-Armed Bandit Problems. *Foundations and Trends in Machine Learning*, 5(1):1–122.
- Dawid, A. (2006). Probability Forecasting. *Encyclopedia of Statistical Sciences*.
- Dawid, A. P. (2007). Fundamentals of Statistical Causality. Research Report 279, Department of Statistical Science, University College London.
- Gifford, A. M., Cohen, Y. E., and Stocker, A. A. (2014). Characterizing the Impact of Category Uncertainty on Human Auditory Categorization Behavior. *PLoS Computational Biology*, 10(7).
- Gopnik, A., Glymour, C., Sobel, D., Schulz, L., Kushnir, T., and Danks, D. (2004). A Theory of Causal Learning in Children: Causal Maps and Bayes Nets. *Psychological Review*, 111:3–32.
- Hagmayer, Y. and Meder, B. (2013). Repeated Causal Decision Making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(1).
- Hagmayer, Y. and Sloman, S. (2009). People conceive of their choices as intervention. *Journal of Experimental Psychology: General*, 138:22–38.
- Meder, B., Hagmayer, Y., and Waldmann, M. R. (2009). The role of learning data in causal reasoning about observations and interventions. *Memory & Cognition*, 37(3):249–264.
- Meltzoff, A. N. (2007). Infant’s causal learning: Intervention, observation, imitation. In Gopnik, A. and Schultz, L., editors, *Causal learning: Psychology, philosophy, and computation*, pages 37–47. Oxford University Press.
- Osman, M. (2010). *Controlling uncertainty: Learning and decision making in complex worlds*. Chichester, United Kingdom: Wiley.
- Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, UK.
- Saito, M. and Shimazaki, T. (2011). Causal Reasoning in Decision Making: The Role of Causal Models and Their Parameters. In Kokonov, B., Karmiloff-Smith, A., and Nersessian, N. J., editors, *European Perspectives on Cognitive Science*.
- Sloman, S. (2005). *Causal Models: How People Think About the World and Its Alternatives*. Oxford University Press.
- Spirtes, P. and Scheines, R. (2001). *Causation, Prediction, and Search, Second Edition*. MIT Press.
- Steyvers, M., Tenenbaum, J. B., Wagenmakers, E.-J., and Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, 27:453–489.